Simulation of the nucleation and growth of biominerals

by

Riccardo Innocenti Malini

Thesis submitted to the University of Sheffield for the degree of

Doctor of Philosophy

Supervised by:

John H. Harding

Colin L. Freeman

May 3, 2016

Contents

3 Simulation of calcium phosphate ions in solution: Analysis of nucleation pathways 57

Appendices

List of Figures

List of Tables

Acknowledgements

I would like to express my sincere appreciation to my supervisors Prof. J. H. Harding and Dr. C. L. Freeman for their guidance, support and comprehension during the past four years. Their incredible knowledge in the area and enthusiasm were essential to fulfil the aims of this investigation. They also helped me to develop both as a person and as a researcher over this time.

I gratefully acknowledge the EPSRC funded consortium, MIB, which allowed me to pursue my PhD. The provided financial support allowed me to meet and collaborate with a large number of people. Especially fruitfull discussion and collaboration were carried with: Dr. A. Finney, Prof. M. P. Rodger, Prof. J. D. Gale and Prof. N. A. J. M. Sommerdijk.

A special thank goes to Shaun A. Hall for his friendship and advices throughout the PhD. The MESAS group has also been invaluable for suggestions and ideas. I would like to especially thank David Sparks, James Miller and Kai Wolff.

Lastly, but also most importantly, I would like to thank my family and friends. My dad and my mum have been a constant source of inspiration and have support throughout the past 26 years. I could never thank them enough. I would also like to thank the friends who have been essential throughout my life and to achieve this goal: Uga didone, Baraibar Sergio, De Saeger Till, Dharana Jawayardene, Sam Matthew and Polytimi Sofotasiou.

Declarations

I declare that all the following work has been carried by the author except in the following cases:

- The force field used in Chapter 3 was obtained from J. D. Gale from Curtin University.
- In Chapter 4 The structures ACC_{Ik} and ACC_{RLJ} were supplied by Y. G. Bushuev from the University of Warwick.
- Many of the information presented in Appendices A and B were calculated by J. D. Gale from Curtin University.

Chapter 4 was mostly extracted from the following publication:

> R. Innocenti Malini, Y. G. Bushuev, S. A. Hall, C. L. Freeman, M. P. Rodger, and J. H. Harding, Using simulation to understand the structure and properties of hydrated amorphous calcium carbonate, CrystEngComm, vol. 18, no. 92, pp. 92- 101, 2016.

Summary

Nature appears to always be a step ahead. Through millions of years of evolution it has developed an incredible array of tools and machinery for a multitude of different applications. Biomineralization is one of those processes. It leads to the formation of a material with functions that include lenses, exoskeletons, grinding tools, mineral reservoirs and shells. All of these materials, are tailor-made for their applications by a complex interplay between the mineral phase and the organic matter. Understanding the fundamental mechanisms leading to the formation of these functional materials would lead to the development of new manufacturing methods to control matter at the nm scale.

Nucleation is a process that is still not well understood. Recent investigation suggests the existence of (meta)stable complexes prior to the precipitation of a wide number of mineral systems. Due to the current resolution limit, it is not possible to directly observe the molecular mechanisms involved. In this thesis, Molecular Dynamics simulations have been used to probe the association of calcium phosphate ionic complexes and their aggregation in solution. To obtain the free energy of the reactions, Umbrella Sampling was used. The results partially confirm the experimental observations showing that the $[Ca(HPO_4)_3]^{4-}$ is stable and can exist in solution. In addition, the aggregation of both

the $[Ca(HPO_4)_3]^{4-}$ and $[Ca(HPO_4)_2]^{2-}$ complexes were analysed allowing for a step-by-step understanding of one of the potential pathways prior to the nucleation of calcium phosphate minerals.

As living organisms have been observed to initially precipitate an amorphous phase prior to the development of crystalline minerals it is important to understand its properties. Hydrous amorphous calcium carbonate is usually the precursor of both calcite and aragonite. As multiple short range structures have been observed for this phase, it is important to sample comprehensively the ensemble of configurations. By studying four models of hydrous ACC obtained in different ways, the structure and role of the water molecules within the structure were analysed. All the structures match experimentally obtained data. Formation of continuous, or percolating, channels was found to be metastable compared to more homogeneous distributions. Additionally the diffusion coefficient of the ions were extremely small suggesting that large scale reorganisation would require a large driving force.

Biomolecules, including proteins, proteoglycans, glycosoamminoglycans, amino acids and peptides have the ability to dramatically alter the nucleation and growth process of biominerals. Often they induce the formation of minerals with non-equilibrium geometries. Here the effect of small organic molecules on the association of ions in solution and on the surface of an amorphous intermediate were investigated using Molecular Dynamics. The results show that organic molecules can stabilise ionic complexes, and can have an important effect on the water dynamics associated with ions.

Chapter 1

Introduction

Biomineralization is the process by which living organisms precipitate minerals that are often essential to their survival. It often results in materials with properties and morphologies tailor-made for specific applications. These could for instance include: protection, motion, storage, optical sensing and cutting. *In vitro* it has not yet been possible to reach a comparable degree of control and thus there is enormous scope for understanding the fundamental mechanisms behind the ability of living systems to manipulate mineral precipitation and growth. In the following work molecular simulations have been used to study the steps preceding crystallization for two mineral systems: calcium phosphates and calcium carbonates.

This chapter will start by introducing the literature relevant to calcium carbonate and calcium phosphate minerals. Then classical nucleation theory and the alternatives will be reviewed with a special focus on the formation of pre-nucleation clusters and other "non-classical" mechanisms. Finally, the effect of organic molecules on the aforementioned process will be detailed.

1.1 Calcium phosphate

Calcium phosphate biominerals have a long standing history. Already in mid Neoproterozoic times (1000-543 millions year ago) records of calcium phosphate biomineralization in eukaryotes have been found in microfossils where apatite was observed to form scales which enveloped the organisms [24]. Interestingly apatite as a biomineral did not become widespread in living organisms due to its importance for their metabolism. By contrast its amorphous precursor has seen much wider application in the living realm, such as phosphate storage (essential for the metabolism and energy production), pressure perception and detoxification [25]. Apatites are mainly found in vertebrate and brachiopods and the reason could be the constant remodelling of the skeleton required by these species during growth [25]. Calcium phosphates minerals are ubiquitous in nature, and can also be found in igneous rocks [26].

In this section the calcium phosphate based minerals relevant to biomineralization will be discussed. The section will start with hydroxyapatite, which is the thermodynamically most stable phase. Octacalcium phosphate and the amorphous precursor will also be introduced as they are important in biomineralization and will be discussed in subsequent chapters. Other phosphate phases such as tricalcium phosphate and monohydrocalciumphosphate have also been observed however they are not as common in living organisms and thus they will not be considered.

1.1.1 Hydroxyapatite

Investigation of this mineral phase is not only encouraged by its presence in our bones and teeth, but also by the wealth of potential applications. Example range from the production of biomimetic biomaterials and synthetically engineered tissues for medical use [27–30], to heavy metal sorption from contaminated soils [31], liquid column chromatography [32] and catalysis [33].

Hydroxyapatite (HAP), chemically defined as $Ca_{10}(PO_4)_6(OH)_2$, can be found with two different crystal structures; a monoclinic and an hexagonal lattice. As the interest here lies predominantly in biologically formed minerals, the hexagonal structure will be discussed. Hexagonal HAP is defined by the $C6_3/m$ space group with $a = b = 9.43$ Å, $c = 6.88$ Å and $\alpha = \beta = 90°$ and $\gamma = 120°$. In Figure 1.1*a* the structure is shown projected down the c axis while in b the structure is projected down the b axis. Within the structure the Ca^{2+} ions occupy two lattice sites as shown in Figure 1.1*a* identified as Ca_1 (coloured in green in Figure 1.1*a*) and Ca_2 (coloured in dark blue in Figure 1.1*a*). The Ca_1 sites are arranged in columns perpendicular to the basal plane while the Ca_2 sites form hexagonal channels, again perpendicular to the basal plane as seen in Figure 1.1*a*. At the centre of these hexagonal channels the hydroxide ions are arranged in a column. The direction of the hydroxide ions within the columns was only recently calculated using density functional theory [34]. De Leeuw showed that in the most stable configuration the hydroxide ions are oriented in opposite directions in neighbouring columns as shown in Figure 1.1*b* [34]. Biological apatite is known not to be stoechiometric and defects such as impurities and vacancies have often been observed. Often the PO_4^{3-} ions are replaced by CO_3^{2-} ions leading to carbonated apatites which are found in vivo [35, 36]. Another important incorporation involves the F^- ion instead of the OH^- ion [37]. This leads to a decreased solubility of the mineral phase which has been exploited in toothpaste and dentistry to help prevent teeth deterioration [38]. Replacement of the Ca^{2+} ions has also been observed. Mg^{2+} can be incorporated instead of the Ca^{2+} in HAP [39]. Due to the aforementioned modifications in biological HAP the Ca/P ratio observed is closer to 1.5 rather than the stoechiometric 1.67.

As already mentioned, HAP is of essential importance due to to its presence in bone and teeth. Bone is an exquisite example of biomineralization. In this case a scaffold composed mainly of collagen is deposited first as shown in Figure 1.2. This is then followed by mineralization along the length of the fibril. Mineral deposition is driven both by interactions between ions in solution and the collagen molecules, but more importantly by non collagenous proteins (NCP) [40–43]. The living system uses various techniques to manipulate the process leading to the hierarchical structure spanning various length-scales shown in Figure 1.2, from the sub-nm scale to the macroscale. The result is a composite material with extraordinary properties. The mineral is mechanically stiff and strong while the fibres help increase the toughness and elasticity of the material which is comparable in strength to steel. Another interesting example of hydroxyapatite within living systems is within kidney stones, as observed using Raman spectroscopy [44].

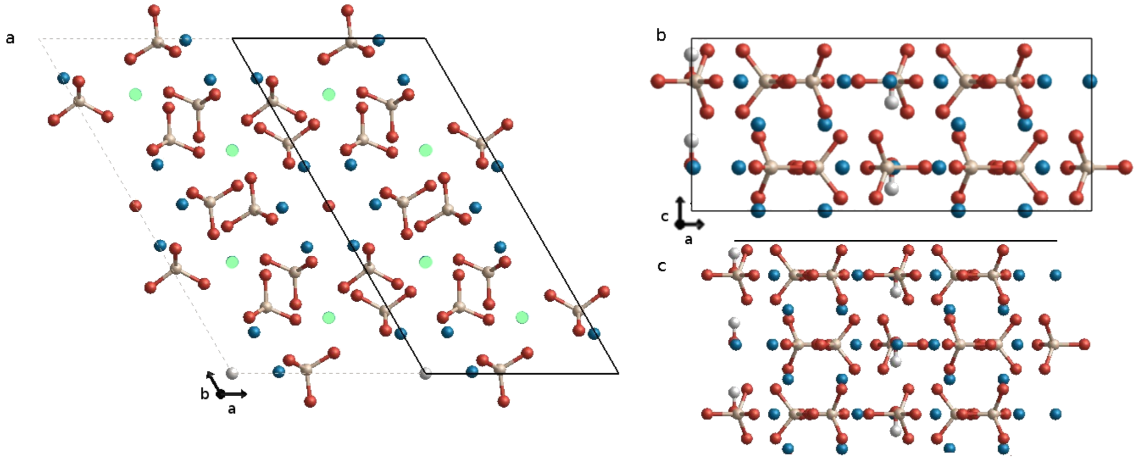


Figure 1.1: (a) HAP unit cell projected down the c and (b) b axis. In the images the oxygens are shown in red, the hydrogens in white, the calciums in blue and the phosphorous in beige. In (a) the Ca^{2+} are coloured in either green or blue depending on their position within the unit cell. Please refer to the text for more information.

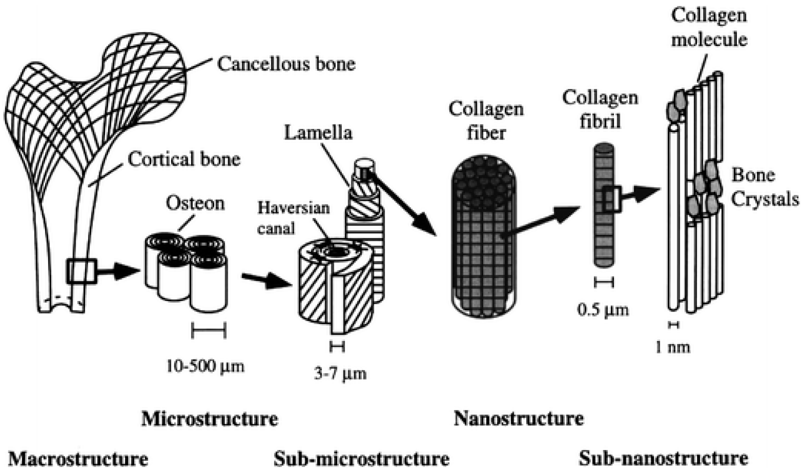


Figure 1.2: Bone hierarchical structure; from the sub-nanostructure to the macrostructure [1]. Reprinted with permission from Elsevier.

1.1.2 Octacalcium phosphate

Brown determined the structure of octacalcium phosphate (OCP),$Ca_8H_2(PO_4)_6(H_2O)_5$, using X-ray diffraction [45]. It is described by the *P -1(2)* space group with crystal lattice parameter a = 19.87 Å, b = 9.63 Å and c = 6.87, and angles α = 89.28, β = 92.22 and γ =108.95 [45]. When observing the structure along the [100] direction, it is easy to see why it could easily be the precursor to hydroxyapatite [46]. This is shown in Figure 1.3 where the hexagonal channels formed by the Ca_2 discussed above form half hexagonal channels around the oxygens of the phosphate. The Ca_1 (in light green in Figure 1.3 are arranged in a column as for HAP. The “hydroxyapatite” like planes are sandwiched by a layer of water and ’non-apatitic’ phosphate ions [45, 47]. The apatitic layers of OCP may act as the nucleation sites for the formation of HAP [46].

In living organisms OCP is rarely observed as the final phase but it is important as a precursor. In vivo it was observed within the gap between the collagen fibril [47–49] before the nucleation of HAP. Crane *et al.* observed it in rat calvarial bone using Raman spectroscopy [48].

OCP shows both osteoinductive and osteoconductive properties leading to potential applications for the manufacture of biomaterials for bone repair [50]. OCP coatings sometime show faster healing rates than HAP or carbonated biological HAP [50–52]. The hypothesis is that the higher solubility of OCP increases the local concentrations of calcium and phosphate ions which would then lead to a chain of chemical signals that elicit a specific reaction in the relevant cells [50].

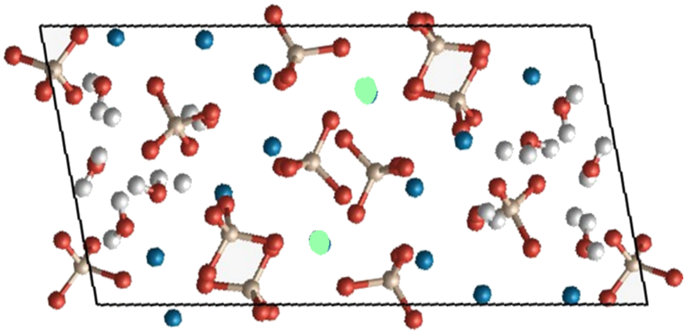


Figure 1.3: OCP unit cell projected down the [100] surface. In the images the oxygens are shown in red, the hydrogens in white, the calciums in blue and the phosphorous in beige. The Ca^{2+} are coloured in either green or blue depending on their position within the unit cell. Please refer to the text for more information.

1.1.3 Amorphous calcium phosphate

As a precursor to both the nucleation of bone, and a potential coating for medical implants, the structure and properties of amorphous calcium carbonate (ACP) have been thoroughly reviewed [53, 54]. ACP is a typical amorphous material with no definite structure and an absence of long range order. The levels of hydration reported lie between 3 and 4.5 water molecules per formula unit and Dorozhkin proposed the existence of a number of configurations with structural features depending on the environment in which it precipitates [54, 55]. To further complicate the process, ACP can be composed of a complex mixture of phosphate groups and can include, PO_4^{3-}, HPO_4^{2-} or $H_2PO_2^{1-}$ ions depending on the pH and the extent of the reaction. The usual calcium to phosphate ratio for this structure reported in the literature lies between 1.1 and 1.6 [54]. However in a recent article published by Habraken *et al.* (2013) a ratio of Ca to P of 0.67 was reported.

It was obtained through pH and Ca^{2+} ion selective electrode (ISE) measurements and subsequent calculation of the chemical composition via mass and charge balance [56]. This was also further confirmed in another independent experiment by using energy-dispersive X-ray spectroscopy (EDX) [57]. Interestingly, after preparation by the usual method of filtration and drying followed by energy dispersive X-ray analysis, they found that their sample reached a value of 1.5 [56], in line with previous experiments. In this case the as-prepared ACP chemical formula was proposed to be $[Ca_2(HPO_4)_3]^{2-}$ [56].

Within living organisms, ACP, as for OCP, is important as a precursor to the final crystalline phase [53,58]. An interesting recent study finally uncovered part of the mineralization mechanism [59]. It was observed that bone-lining cells, osteoblasts, concentrate intracellular vesicles containing mineral granules close to the membrane. Elemental analysis and electron diffraction showed that the mineral granules were amorphous and, surprisingly, that the Ca/P ratio was 0.75 [59]. This is very similar to the recent observations obtained from *in vitro* experiments [56, 57]. ACP particles have also been found in mammalian milk where they are a highly soluble reservoir of calcium and phosphate ions ready to be absorbed by the body.

1.2 Calcium carbonate

Eukaryotic biomineralization of calcium carbonate also started in mid-Neoproterozoic times. However in contrast to calcium phosphate, calcium carbonate skeletons quickly became widespread during the Cambrian explosion (began about 540 MYr ago) and can be observed in all types of eukaryotic organisms; plants, protists and animals [25].

Calcium carbonate is a versatile mineral system and can be found in different crystalline forms, called polymorphs. For this mineral, both anhydrous and hydrated crystalline structures exist. The anhydrous polymorphs in order of stability are calcite, aragonite and vaterite. Monohydrocalcite and ikaite are the hydrated crystalline polymorphs. Another extremely important form of this mineral system is amorphous calcium carbonate (ACC), which precipitates before the crystalline phases in many cases during biomineralization. It is believed to be an important part of the process by which living organisms control the final morphology and properties of the crystalline material [12]. All of these phases will be briefly reviewed in the following discussion.

1.2.1 Calcite

This crystalline phase of the calcium carbonate system is ubiquitous in nature as it is the thermodynamically most stable phase. Calcite has a rhombohedral crystal structure with space group $R-3cH$. The unit cell dimensions, as obtained from X-ray diffraction by Maslen *et al.* (1993) are $a = b = 4.991$ Å, $c = 17.062$ Å, $\alpha = \beta = 90.0°$, and $\gamma = 120.0°$ [60]. In Figure 1.4a the unit cell is shown along the c and a directions. It can be seen that the Ca^{2+} and the CO_3^{2-} ions are arranged in alternative layers stacked one on top of the other perpendicular to the c axis. The CO_3^{2-} ions in alternate layers point in different directions as shown in Figure 1.4a and 1.4b. Within their respective planes all the ions occupy an hexagonal lattice. In the bulk mineral the Ca^{2+} ions has six oxygens in its coordination shell and six carbon atoms from the CO_3^{2-} ions in its coordination shell (Table 4.5). This implies that the Ca^{2+} has a monodentate interaction with the CO_3^{2-}

ion. Calcite has a density of 2.72 g.cm^{-3} and a hardness of 3 on the mohr scale. When calcite is precipitated from solution without additives, the structure obtained shows a typical faceted morphology arising from the underlying structure of the unit cell and expresses the most stable (10 $\bar{1}$4) surface shown in Figure 1.4*c*. Other surfaces have also been observed when calcite is precipitated in the presence of additional molecules or in non-standard conditions (e.g. constrained volume).

Calcite is found in many living species. A typical example is birds' eggshells. These need to provide protection, have the ability to exchange gases and liquids with the surrounding environment, and act as a calcite reservoir "which will be consumed by the growing embryo during ossification" [61]. Its properties are essential for the growth of the bird embryo and are therefore tightly controlled [62]. The shell is composed of 96% calcite crystals with specific orientations and thicknesses depending on the species. These material parameters correlate strongly with genetic variation and thus with the organic molecules present in the eggshell [62]. They play an essential part in the control of both the location of crystal nucleation and in the growth of the crystal [63]. Another example of calcite in nature are the coccospheres, the shells of coccolithophores, which make up a large part of our beaches. In Figure 1.5 a set of time lapse photographs of the expulsion of a calcite coccolith is shown. The crystal grows within vesicles in the single cell organism as shown in the first two images of Figure 1.5. Within this enclosed environment, the solution chemistry, the presence of organic molecules and the physical constrains can be tightly controlled leading to the formation of a single crystal with a morphology tailor-made at the nm scale. The

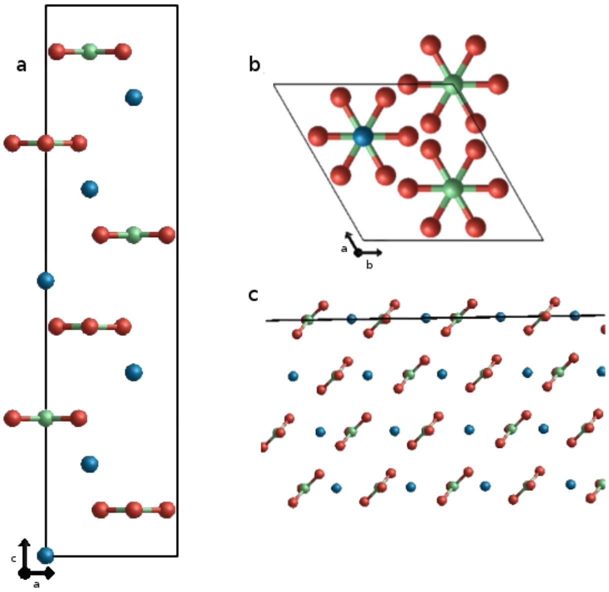


Figure 1.4: Calcite structure projected down the a) c a plane, b) b a. c) 104 plane of calcite. In the images the oxygens are shown in red, the calciums in blue and the carbons in green.

process ends with the expulsion of the shell through exocytosis as seen in the third image in Figure 1.5 [41].

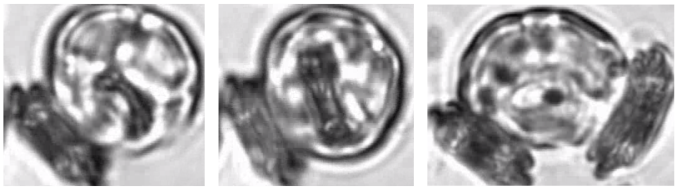


Figure 1.5: Time lapse photographs of coccolith formation by a coccolithophore.

Table 1.1: Coordination of the calcium ion in the different polymorphs of the calcium carbonate system. In the table, oxygens in general are named O_{all}, O_W is used for the waters' oxygens and O_C for the carbonates' oxygens.

System	C_N(Ca-O_X)			C_N(Ca - C)
	O_{all}	O_W	O_C	C
Calcite	6	0	6	6
Aragonite	9	0	9	6
Vaterite	8	0	8	6
Monohydrocalcite	8	2	6	4
Ikaite	8	6	2	1
ACC	5-9, 7.4 ± 0.5	-	-	4.5 ± 2

1.2.2 Aragonite

Often observed in the shells of molluscs, aragonite is another polymorph of the calcium carbonate system encountered in living organisms. Experiments show that aragonite is less stable compared to calcite by $\Delta^{calcite}_{aragonite}$G = -840±20 J.$mol^{-1}$ where $\Delta^{calcite}_{aragonite}$H = 440±50 J.$mol^{-1}$ [64]. Aragonite has an orthorhombic crystal structure, with space group *Pmcn*. The experimental structure obtained by De Villiers (1971) had a unit cell with a = 4.9614Å, b = 7.9671Å and c = 5.7404 Å, and $\alpha = \beta = \gamma = 120°$ [65]. From Figure 1.6a it can clearly be seen that unlike calcite, in aragonite the CO_3^{2-} ions are arranged in a double layer. This double layer of carbonate ions is sandwiched by layers of Ca^{2+} ions. The Ca^{2+} layers are twice as dense as the carbonate layers to balance the electric charge. Another difference between calcite and aragonite lies in the coordination shell of the Ca^{2+} ions. The calcium ions are surrounded by nine oxygen atoms coming from six carbonate ions. This entails that three of the CO_3^{2-} ions interact with the Ca^{2+} ions via a bidentate interaction (See Table 4.5). The more compact organisation around the calcium ions is why this structure is denser than calcite, with an experimentally measured value of 2.93 $g.cm^{-3}$.

Aragonite precipitated from solution has a much more interesting morphology than calcite and different surfaces are expressed as shown in Figure 1.6c. This stems from a much smaller difference between the surface energies of the different surfaces of this mineral than for calcite, as shown in Table 1.2 [2]. For comparison, the three most stable surfaces of calcite are given in Table 1.2.

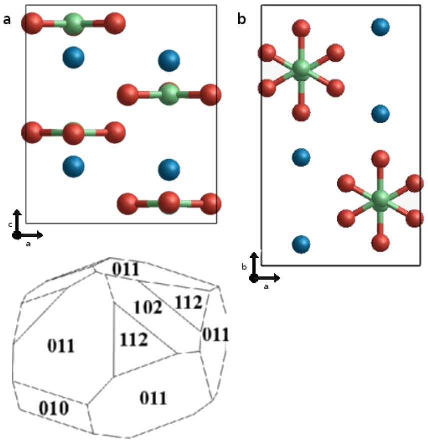


Figure 1.6: Aragonite structure projected down the a) c a plane, b) b a plane. c) Example of nanoparticle morphology obtained using the Wulff method [2]. In the images the oxygens are shown in red, the calciums in blue and the carbons in green.

In living organisms, aragonite based materials are often found to have an exoskeletal function. They are also found in molluscs, where the shell is formed of alternating layers of calcite and aragonite separated by an extracellular matrix (ECM). The ECM is composed of β-chitin and it is an important element for the correct polymorph selection [66]. The most interesting feature of this structure lies in the spatial segregation of the calcitic and aragonitic minerals. Living organisms achieve it by the precise placing of macromolecules, within the extracellular matrix, with the ability to preferentially nucleate one of the two polymorphs [66–68]. The presence of macromolecules able to select a specific polymorph is essential for the precipitation of aragonite, since under standard conditions calcite is thermodynamically the most stable polymorph.

Table 1.2: Aragonite and calcite surfaces from the most stable to the least as detailed by Bano *et al.* [2]

Aragonite surface	Surface energy ($J.m^{-2}$)	Calcite surface	Surface energy ($J.m^{-2}$)
(011)	0.8406	(104)	0.7112
(102)	0.8780	(105)	0.8125
(112)	0.8955	(103)	0.9845
(001)	1.0161		
(110)	1.0350		
(012)	1.0848		
(221)	1.1041		

1.2.3 Vaterite

The structure of this biomineral has been a subject of debate for many years. Both experimental and theoretical investigations have suggested several space groups that could describe the structure: *Pbnm* [69,70], $P6_3/mmc$ [71], *Ama2* but with misaligned domains offset by 120 ° [72], $P6_522$ [73] and $P3_221$ [74]. One of the main issues in determining the structure of vaterite is the disorder associated with the carbonate ion. This was shown by Wehrmeister *et al.* [75]. While investigating the structure using Raman spectroscopy, they identified three unequivalent carbonate ions which could not be reconciled with any of the models proposed [75]. Demichelis *et al.* used first principle calculations to show that the different structures of vaterite were separated by barriers of the order of kT at room temperature [74]. This suggests that vaterite could be composed of a number of different structures. Additional investigations, both theoretical and experimental, concluded that the structure is characterised by a number of unit cells stacked upon each other [76–79]. The most important cells are described by the following space groups: $P3_221$ and its mirror image $P3_121$, *C1* and *C2* [76].

Vaterite is the least stable form of anhydrous calcium carbonate mineral. Thus it is rarely seen as an end product of biomineralisation. However, it has been observed as an intermediate phase on the pathway to the formation of both calcite and aragonite [80, 81]. One rare example of its presence in living organisms is stabilised by biomolecules in the shell of the freshwater snail *Biomphalaria glabrata* [10].

1.2.4 Hydrated mineral phases

Hydrated mineral phases of calcium carbonate are also found in specific geological environments and can be synthesised in *vitro* in the presence of Mg^{2+} ions. Monohydrocalcite contains one water molecule per formula unit, $CaCO_3.H_2O$. It has an hexagonal crystal structure and is described by the *P31* space group [82]. The calcium ion has eight oxygens in its coordination shell. Six of them arise through the interaction with four carbonate groups with two of them being involved in a bidentate interaction. The other two oxygens come from two water molecules as shown in Table 4.5.

Ikaite is much more hydrated, containing six water molecules per formula unit, $CaCO_3.6H_2O$. The lattice is monoclinic (space group $C12/c1$) with dimensions determined as $a = 8.8268$ Å, $b = 8.3235$ Å, $c = 11.0532$ Å, $\alpha = \gamma = 90$ ° and $\beta = 110.579$ ° [83].

As far as I know, neither of these polymorphs of calcium carbonate is produced by controlled biomineralization within living organisms.

1.2.5 Amorphous calcium carbonate (ACC)

ACC is truly amorphous in the sense that pair distribution functions obtained from diffraction experiments show the absence of order above 10 Å [13, 84, 85]. However, both short (first coordination shell) and medium range (within 10 Å) order can be observed and experiments by Gebauer *et al.* showed that ACC precipitated at different pHs showed structural features pertaining to different polymorphs [55, 86]. The less stable ACC obtained at lower pH was similar to vaterite, while the more stable ACC was more similar to calcite [86]. The coordination number of the Ca^{2+} ions varies between five and nine, further demonstrating the potential for multiple amorphous structure of this metastable phase [10,11,13–16]. The distribution encompass all the different crystalline phase of calcium carbonate as shown in Table 4.5. ACC can be found with different compositions (mostly changes in the water content). The hydrous phase is more stable than its anhydrous counterpart which quickly transform to one of the crystalline polymorphs when in solution at room temperature [11,12]. Thus water could stabilise the structure. Using molecular dynamics simulations, it was found that up to a size of approximately 3 nm, the hydrous phase was more stable than a calcite nanoparticle containing the same number of formula units [87].

In living organisms ACC is ubiquitous. It is found at the growth front of minerals as a precursor, but it is also important as a structural material and as an accessible reservoir of ions [41]. In living organisms it acts as a precursor during the formation of mollusc shells, sea urchin larval spicules and in nacre as a precursor to aragonite [88–90]. It is believed to be both more stable than calcite at the nanoscale and to be an essential step in the formation of non-equilibrium morphologies [87,91].

1.3 Nucleation and crystallization of biominerals

Nucleation is the process that describes the formation of a secondary thermodynamic phase from an homogeneous phase. It could entail precipitation of a secondary phase from a melt, a gas or a solution.

1.3.1 Classical nucleation theory

Classical nucleation theory (CNT) has now been used for more than a hundred years to quantitatively describe the process leading to the precipitation of a secondary thermodynamic phase from a parent homogeneous phase. Gibbs formulated the theory to describe the nucleation of droplets from vapour [92, 93]. Classical nucleation theory was then developed by a number of researchers including Volmer and Weber, and Becker and Doring [94, 95]. CNT describes the change in free energy as a function of the size of the nascent nucleus. This depends upon two processes. One is the favorable interactions between the monomers in solution. This lead to a gain in free energy and will be here labelled G_b. The other process is the unfavourable formation of an interface with the parent solution, thus an interfacial tension will exist between the nascent nuclei and the parent solution, which energy is here labelled as G_s. The change in free energy as a function of the number of molecules in the nucleus is then the addition of the free energy change due to the bulk (ΔG_b) and surface term (ΔG_s):

$$\Delta G = \Delta G_b + \Delta G_s \tag{1.1}$$

where (ΔG_b) and (ΔG_s) are defined by:

$$\Delta G_b = \left(\frac{\frac{4}{3}\pi r^3}{\Omega} \right) \Delta\mu \tag{1.2}$$

$$\Delta G_s = 4\pi r^2 \alpha \tag{1.3}$$

In the two equations above, r is the size of the nuclei in Å, Ω is the volume per molecule, α is the interfacial energy and $\Delta\mu$ is the change in chemical potential defined as:

$$\Delta\mu = K_B T ln\sigma \tag{1.4}$$

where K_B is Boltzmann's constant, T is the temperature and σ is the supersaturation and is equivalent to:

$$\sigma \equiv ln\frac{AP}{K_{sp}} \tag{1.5}$$

where K_{sp} is the equilibrium constant and AP is the activity product of the solution. The change in free energy, ΔG, as a function of the nucleus size is plotted in Figure 1.7 together with G_s (green) and G_b (red). Since G_b and G_s scale as the cube and the square of the radius respectively, the change in free energy as a function of the size of the nuclei will pass through a maximum (ΔG_{ex}) at a certain critical radius r_{crit}. Clearly the growing nuclei will need to overcome this barrier to be able to grow. Thus a nucleus that form in the parent solution of size below r_{crit} will tend to dissolve. Large enough stochastic fluctuations in the solution could lead to a nuclei with a size equivalent to r_{crit} or larger, which will then be able to grow. The critical radius, r_{crit} can be obtained via the following relationship:

$$\frac{\delta \Delta G}{\delta r} = \frac{4\pi r^2}{\Omega}\Delta\mu + 8\pi r\alpha \tag{1.6}$$

At the maximum, where $\frac{\delta \Delta G}{\delta r} = 0$:

$$r_{crit} = \frac{-2\alpha\Omega}{\Delta\mu} \tag{1.7}$$

r_{crit} can be inserted back into equation 1.1 to obtain ΔG_{ex}:

$$\Delta G_{ex} = \frac{\frac{16}{3}\pi\alpha^3\Omega^2}{\Delta\mu^2} \propto \frac{\alpha^3}{\sigma^2} \tag{1.8}$$

Although the original formulation of CNT did not consider the precipitation of metastable states, such a process would not depart from classical nucleation theory. This was first observed by Ostwald in 1897 and was called the “step rule”. It states that the first phase that precipitates does not have to be the thermodynamic most stable phase. Instead phases that have a structure, and a free energy closer to that of the parent phase could precipitate first as shown in Figure 1.8. The metastable precursors would then transform directly, or with additional steps, into the thermodynamically most stable phase. In Ostwald’s time

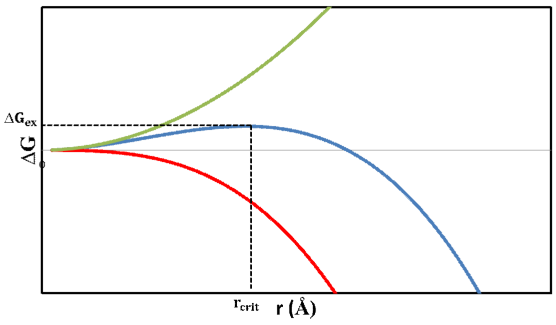


Figure 1.7: Change in Gibbs free energy as a function of nucleus size. The total change in Gibbs free energy as a function of the nucleus radius is shown in blue, the change in G_s in green while the red curve represents the change in G_b.

it was not possible to demonstrate the mechanism. The rule of stages was subsequently demonstrated for large units, such as proteins and colloids, both experimentally [96,97] and theoretically [98, 99]. The proof for inorganic materials was only obtained recently, when Chung *et al.* used high resolution electron microscopy to study the nucleation of $LiFePO_4$. They observed the precipitation of different metastable phases before the formation of the final crystalline structure [100]. The formation of transient states has also now been observed in many cases in biomineralisation. This does not refer only to the initial formation of an amorphous precursor [15,53,58,101] but also to polymorphs that are metastable with respect to the most stable crystalline phase [80, 81]. Such a process has been observed in the calcium carbonate system where precipitation following the scheme; hydrous ACC → anhydrous ACC → vaterite → calcite has been reported many times [80, 81]. Crystallisation thus would occur through an initial increase of the local concentration of ions forming a disordered initial precipitate. Then the crystalline final polymorph could form via solid

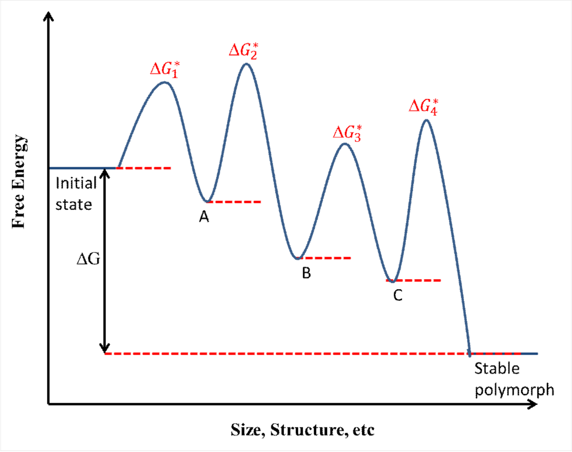


Figure 1.8: Ostwald's step rule. Multiple metastable states are present in the free energy landscape each associated to an energy barrier distinguished by *. The total change in free energy is ΔG

.

transformation or via a dissolution reprecipitation mechanism.

CNT gives an excellent qualitative account of the process of nucleation but has serious shortcomings in quantitative predictions [97]. The shortcomings in the quantitative descriptions stems from the assumptions on which CNT is based. One of the most debated is the "capillary assumption", which assumes that the interfacial tension and the cohesive energy are identical to the macroscopic values. It can also be seen from equation 1.1 that CNT assumes that the nucleus is a sphere and that its properties are isotropic.

1.3.2 Alternatives to classical nucleation theory

The most interesting observation regarding the nucleation of biominerals is the formation of clusters of ions in the calcium carbonate system before the nucleation event [3].

The formation of this species, could have an important impact on the additional pathways by which a living organism can manipulate the process. By using titration of Ca^{2+} ions in a carbonate containing solution, it was observed that a fraction of the Ca^{2+} was bound in solution and that it depends linearly on its total concentration in solution [3, 102]. Since titration with CO_3^{2-} gave the same result, it was concluded that Ca^{2+} and CO_3^{2-} ions were bound in a neutral complex where the $Ca{:}CO_3$ ratio is 1:1 [3, 102]. Analysing the same process at different pHs showed that the fraction of Ca^{2+} ions bound in the clusters depended on the concentration of CO_3^{2-} present in the solution. As this implies thermodynamic equilibrium the energy barrier involved in the binding of the ions must be below room temperature energy. The binding of Ca^{2+} and CO_3^{2-} ions could be the formation of ion pairs. However, further investigation using Analytical Ultracentrifugation (AUC) showed that there were two different types of solutes in the solution, one with a size corresponding to the ion pair and the other one larger, $\approx$ 2 nm or 70 ions as shown in Figure 1.9 [3, 102].

The observation was further confirmed by cryo-TEM [103]. In this case the ions were slightly smaller, between 0.6 and 1.1 nm, mainly due to the inclusion of the hydration shell in the AUC experiments. They speculated that the clusters would then aggregate to form ACC. Additional clues were obtained from Molecular Dynamics simulations. Demichelis *et al.* showed that the addition of an ion pair to the cluster was favourable with a $k >> 1$ and that they formed a dynamic equilibrium with the ions in solution [20]. The clusters showed various configurations including linear, branched and circular. In addition, ion attachment and detachment was observed. The simulated clusters were initially obtained in a solution

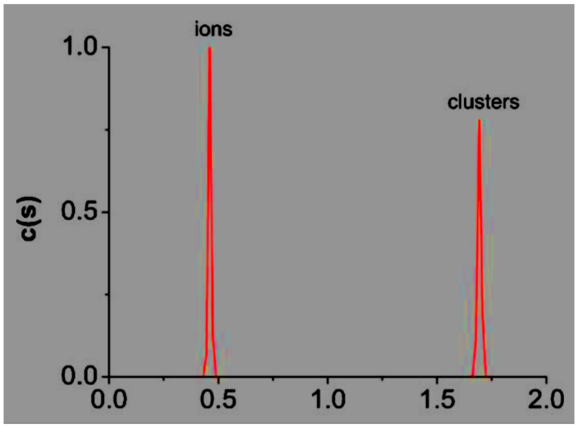


Figure 1.9: Distribution of sedimentation coefficients at pH 9 for a solution of Ca^{2+} and CO_3^{2-} ions. The distribution clearly show the presence of larger species in addition to the ion pair [3].

at much higher concentration than experiments and were then inserted in solution of much lower concentration. A recent investigation by Finney showed that if the simulation at the lower concentration is performed for longer timescales the clusters break up, leading to questions about their stability and the role of kinetics in the work of Demichelis *et al.* [20]. Importantly, in all investigations it was shown that the clusters were highly hydrated and thus had to be considered as solute.

A recent review by Gebauer *et al.* [104] proposed the subsequent definition for a pre-nucleation cluster (PNC):

1. PNCs are composed of the constituent atoms, molecules, or ions of a forming solid, but can also contain additional chemical species.
2. PNCs are small, thermodynamically stable solutes, and there is thus formally no phase boundary between the clusters and the surrounding solution.

3. PNCs are molecular precursors to the phase nucleating from solution, and hence participate in the process of phase separation.

4. PNCs are highly dynamic entities, and change configuration on timescales typical for molecular rearrangements in solution (i.e., within hundreds of picoseconds).

5. PNCs can encode structural motifs resembling, or relating to, one of the corresponding crystalline polymorphs.

In their review, they suggested that growth through aggregation would not be possible for pre-nucleation clusters. The rationale is that they do not form an interface with the solution and thus have no interfacial energy. Hence there is no driving force towards aggregation (release of excess energy). Instead, following the recent simulation from Wallace *et al.*, the clusters could keep on growing down a monotonically decreasing energetic landscape [105]. As the clusters grow, the coordination number of the ions increases from two to three suggesting the formation of a more condensed phase [105]. It suggests that the clusters could have a range of sizes depending on the supersaturation and their growth would lead to a liquid-liquid separation as was first suggested by Faatz *et al.* [105,106].

There is still disagreement on this mechanism. In 2012, Hu *et al.* analysed the nucleation of calcium carbonate solutions on self-assembled monolayers (SAMs) [107]. Their investigation included a theoretical investigation of classical nucleation theory. By assuming that the interfacial energy was not constant but varied as a function of the ion association size, they showed that ΔG_{ex} would be modified accordingly:

$$\Delta G_{ex} \propto \frac{\alpha^3}{(\sigma \pm C)^2} \tag{1.9}$$

where C is a geometrical factor and depends on the shape of the cluster, its radius but also the stability of the clusters with respect to the ions in solution or ion pairs. If the cluster is metastable, the + sign applies and the free energy barrier is reduced. If the cluster is stable then the minus sign should be used and the free energy barrier to nucleation is increased [56, 104, 107, 108]. They argue that pre-nucleation clusters could not lie on the nucleation pathway as they would increase the nucleation barrier. In a recent study analysing the nucleation of calcium phosphate minerals, Habraken *et al.* used the same treatment to explain the precipitation of amorphous calcium phosphate (ACP) under conditions where using CNT leads to ΔG_{ex} values which were unphysical (CNT predicted that the free energy barrier for the precipitation of ACP was 600 times higher compared to hydroxyapatite) [56]. In contrast to the calcium carbonate system, it was determined that the $[Ca(PO_4)_3]^{4-}$ clusters observed in their study did form an interface with the solution, and thus were likely to be metastable. Hence their aggregation leads to a release of free energy through the decrease of the interface area [56, 109]. For the case of calcium carbonate, a recent computational study, which was run at concentration levels similar to experiment did not lead to the formation of pre-nucleation clusters. Even when clusters were inserted in the solution, these would fall apart at long timescales leading to an equilibrium dominated by ion pairs and hydrated ions (Finney - Goldschmidt 2015). The results observed by Demichelis *et al.* could be due to slow kinetics [20].

1.3.3 Effects of organic molecules on the nucleation of biominerals

In vivo, the ions are often in a solution with multiple additives. Some of these molecules have already been isolated and their effect often characterized using *in vitro* experiments with single molecules to study their direct effect [40] or *in vivo* by gene knockout investigations [110] to try and analyse the role of the protein in the chain of events. In this section the most relevant literature on the subject will be briefly reviewed.

One of the most interesting results observed *in vitro* is the formation of the polymer induced liquid phase (PILP) first observed by Gower and coworkers [91, 111]. The formation of the phase arise from the interaction with polyaspartic acid, an analogue of proteins often used *in vitro*. This organic molecule can stabilise a dense ionic liquid with intriguing properties. By using this process, it was possible to form thin crystalline films on surfaces through a pseudomorphic precipitation which allowed retention of the morphology of the precursor observed in biomineralisation. Another property was capillary action, i.e. it could penetrate holes and crevices due to surface tension and adhesive forces [42, 43, 112]. In the work of Jee *et al.* it was also observed that the molecular weight of the polymer was extremely important in controlling the properties of the forming liquid phase. To stabilize a calcium phosphate PILP, polyaspartic acid of higher molecular weight was needed than for the calcium carbonate system [43]. This was probably the reason behind the lack of capillary action in the calcium phosphate system in the study led by Cantaert *et al* [113]. The formation of the PILP phase did not seem to depend upon the basic or acidic character of the organic molecule. Early studies focussed on acidic molecules due to observation of a large number of proteins rich in glutamic and aspartic acids in molluscs [114], but Cantaert *et al.* showed that positively charged additives like poly(allylamine hydrochloride) could have a very similar effect [115].

Chapter 2

Molecular simulations

2.1 Introduction

The use of computer simulation to obtain a fundamental understanding of the molecular and atomic interactions leading to experimentally observed behaviours, has seen a dramatic increase since its infancy. The incredible development in computer technology has led people to study systems of increasing complexity and size for longer timescales. The result is that *in silico* experiments can start to investigate in a more realistic way the observations and hypotheses resulting from *in vitro* and *in vivo* analysis. Applications of molecular simulations include the calculation of physical properties of nanomaterials [116] and chemical/physical processes in a number of areas including pharmaceutical, renewable energies and reaction catalysis.

All atomistic simulations take advantage of the Born-Oppenheimer approximation which states that the nuclear and electronic motions are decorrelated. This is based on the difference in mass between the nucleus and the electrons of an element. Electrons will be moving at a much higher speed compared to the nucleus and thus will relax to their equilibrium position instantaneously compared to the timescale of nuclear motion [117].

First principles or *ab initio* calculations, treat the electronic interactions and distributions explicitly at each timestep by solving the Schrodinger equation using a number of approximations. The result is an accurate prediction of the energies and configurations of the system under study. However, because of the cost of solving the wave equation for each atom at every timestep, this method is constrained to small systems and short timescales. In contrast, atomistic simulations assume that the interactions between atoms caused by the behaviour of the electrons can be approximated by a function dependent only on the distances between the atoms of interest. This permits a large reduction in the computational cost. An additional reduction in computational cost can be obtained by grouping together selected atoms as single interacting units, for instance in coarse grained simulations. Mixed simulations have also been performed where the level of complexity included in specific regions of the system are dependent upon its importance for the simulation. For instance during *in silico* studies of enzymes, the active site is generally modelled using first principles methods and the surrounding sub-system interactions are calculated via an empirical force field or a continuum model [118].

The subsequent chapter will focus on Molecular Dynamics. In addition methods that allow a more comprehensive sampling along reaction coordinates will also be explored. However, we will start by describing the basic statistical mechanics principles that allow for the calculation of average properties of the system from the position and momentum of the different elements.

2.2 Statistical mechanics

Statistical mechanics provides a way to relate the microscopical properties of a system to its bulk properties [119]. The presented literature is based upon the treatment outlined by Leach, [117] and Frenkel & Smith [120].

2.2.1 Ensembles

A system state can be fully defined by knowing all the coordinates of the particles, q^N, and their momenta, p^N. All the states that can be accessed by the system make up the phase space, where each point corresponds to a microstate. The position and momenta of the elements within a microstate define the Hamiltonian, H, of the system. It corresponds to the sum of the kinetic and potential contributions arising from the interactions between the elements. In statistical mechanics the averages are taken over an ensemble of microstates. $f(q^N, p^N; t)$, which is defined over the system's phase space, gives the relative probability of the different microstates of the system at a time, t:

$$\iint f(q^N, p^N; t) dq^N dp^N = 1 \tag{2.1}$$

This function is dependent upon the constrains applied to the system and the following sections will briefly detail the most important ensembles.

Microcanonical ensemble

It is defined by a constant number of molecules/elements, N, a constant volume V and a constant energy E. This ensemble is not used extensively in the literature as it is rare to see experiments that are performed at constant energy, making it difficult to verify results. An interesting property of this ensemble is that with a fixed energy when the system reach equilibrium the entropy is maximised. The density function of the microcanonical ensemble, f(q^N,p^N), is defined as:

$$f(q^N, p^N) = \frac{1}{W(N,V,E)}\delta[(H(q^N, p^N) - E)] \tag{2.2}$$

Here $H(q^N, p^N)$ is the Hamiltonian of the system, and $W(N,V,E)$ is the sum over all accessible states in the phase space, also called the partition function (a normalization factor). This is essential for equation 2.1 to hold:

$$W(N,V,E) = \iint \delta[H(q^N, p^N) - E]dq^N dp^N \tag{2.3}$$

Canonical ensemble

Instead of the energy, the temperature, is kept constant. Thus the ensemble will have a constant number of particles, volume and temperature, or for brevity NVT. This is the most common ensemble because of the balance between computational cost and a realistic comparison to experiments. In this case the probability density of a state is obtained from:

$$f(q^N, p^N) = \frac{1}{W(N,V,T)}\exp\left(-\beta H(q^N, p^N)\right) \tag{2.4}$$

where β is the inverse product of the Boltzmann factor, k_B, and the thermodynamic temperature, T (Kelvin scale). The partition function, $W(N,V,T)$, for this ensemble is defined by:

$$W(N,V,T) = \iint \exp\left(-\beta H(q^N, p^N)\right) dp^N dq^N \tag{2.5}$$

Isobaric ensemble

This is the thermodynamic ensemble that is most similar to the large part of processes happening *in vivo* and *in vitro*. It is characterized by a constant number of particles, pressure and temperature or in short notation NPT. In most processes, both natural and synthetic, not only the energy but also the volume fluctuates due to microscopic changes in the instantaneous density of the system. The probability density of each state is given by:

$$f(q^N, p^N) = \frac{1}{W(N,P,T)} \exp\left(-\beta[H(q^N, p^N) + PV]\right) \tag{2.6}$$

where the partition function, $W(N,P,T)$, is obtained through:

$$W(N,P,T) = \int_0^\infty \exp\left(-\beta PV\right) W(N,V,T) dV \tag{2.7}$$

Additional ensembles do exist. For instance the Gibb's ensemble, where the chemical potential is kept constant rather than the number of elements [117] or the Blue Moon ensemble which was created to obtain the potential of mean force from unbiased constrained simulations [121]. These, however will not be used in the subsequent investigations and thus will not be detailed here.

2.2.2 The ergodic hypotheses

One of the underpinning factors of statistical mechanics is the ergodic assumption. This states that by sampling all the infinite number of microstates available to the system, the ensemble average of a property will be equal to the macroscopic time averaged quantity. The ensemble average for a property $\langle B \rangle_{ens}$ is defined as:

$$\langle B \rangle_{ens} = \iint B(q^N, p^N) f(q^N, p^N) dq^N dp^N \tag{2.8}$$

where $B(q^N, p^N)$ is the value of the function at specific p^N and q^N, and $f(q^N, p^N)$ is the ensemble specific probability density function. The time average is defined as:

$$\langle B \rangle_t = lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau B(p^N(t), q^N(t)) dt \tag{2.9}$$

Thus for sufficiently long trajectories where all microstates are sampled proportionally to their weight, the value obtained by an ensemble average should equal a time average. This is not usually achievable and thus an ensemble average is almost always an approximation.

2.3 Molecular Dynamics simulations

As mentioned in the introduction the main approximation in classical Molecular Dynamics (MD) simulations is the electron-electron interactions, which are described by functions dependent only on the distance between elements and are constant during the simulation. These interactions are used to obtain the potential energy of the system, $U(q^n)$, which is used in combination with Newton's laws of motion to drive the system along a trajectory. This time evolution of the system can then be used to extract average properties.

2.3.1 Classical potentials

In Molecular Dynamics the force acting upon the particles is obtained by partially differentiating the potential energy, $U(q^N)$ acting on the elements along all three cartesian space vectors:

$$F_i = \nabla_{r_i} U(q^N) \tag{2.10}$$

where the configurational energy, $U(q^N)$, is the sum over all the different contributions arising from the interactions between the atoms, molecules and, if present, external fields [120, 122]. Force-fields for specific systems are usually obtained through fitting to structural parameters (i.e. lattice parameter in a crystal structure obtained via X-ray), physical properties (i.e. solvation free energy), *ab initio* simulations, etc. In the subsequent sections the intermolecular and intramolecular contributions to the potential energy will be presented.

Intramolecular interactions

Intramolecular interactions are necessary to maintain the natural geometry of molecules. In the subsequent sections we will detail the different intramolecular interactions and define the functional forms used during the following investigation.

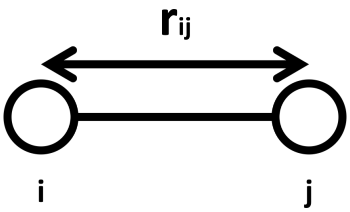


Figure 2.1: Image showing two elements, i and j separated by a distance r_{ij}.

Two body term

This is used to control the distance between two atoms around an equilibrium position and can be used as a representation of a covalent bond. It is usually defined by a harmonic potential function:

$$U(r_{ij}) = \frac{k}{2}(r_{ij} - r_0)^2 \tag{2.11}$$

where r_0 is the equilibrium distance between two atoms within a molecule while r_{ij} is the current distance between element i and j as shown in Figure 2.1. k controls the stiffness of the bond or, how steeply the function increases as the distance diverge from r_0. It is important to note that equation 2.1 tends to infinity as r_{ij} increases. Thus the bond is unbreakable, and the physical behaviour of the interaction is well represented only close to equilibrium. More sophisticated potential functions have been developed to study systems in which bond breaking is allowed. These are used in reactive force fields. However they will not be used in this work and thus will not be discussed.

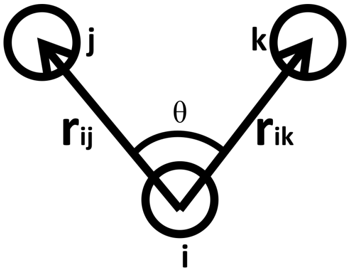


Figure 2.2: Angle θ_{ijl} formed by three atoms, i, j and l.

Three body term

The three body term potential function is added to control the angle between a set of three atoms within a molecule. Like the two body term, a harmonic function is used to define this interaction:

$$U(\theta_{jik}) = \frac{1}{2}l(\theta_{jik} - \theta_0)^2 \tag{2.12}$$

Here θ_0 is the equilibrium angle, θ_{jik} is the current angle formed by the elements i, j and k (Figure 2.2), and l control how easy it is for the angle to depart from its equilibrium position.

Four body term

A four body term is usually inserted to control the dihedral angle composed by a set of four elements within a molecule. In this study a cosine form has been used:

$$U(\phi_{ijkn}) = A[1 + \cos(m\phi_{ijkn} - \delta)] \tag{2.13}$$

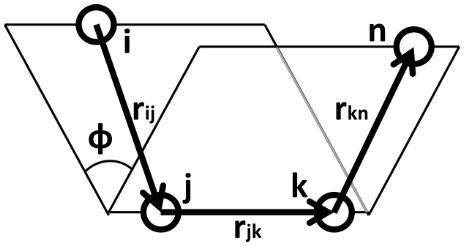


Figure 2.3: Dihedral angle arising from the intersection of the plane passing through atom j, k and n, and the plane defined by atoms i, j and k.

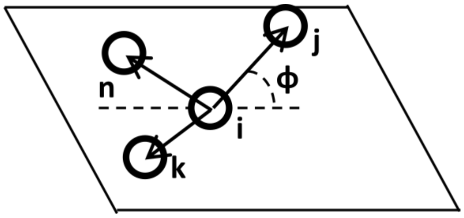


Figure 2.4: Inversion angle formed between the plane composed of atoms n, i and k, and the vector that connects atoms j and i.

where ϕ_{ijkm} is the angle between two planes, one defined by the elements i, j and k, and the other by the elements j, k and n(Figure 2.3).

An additional four body term was used to simulate the carbonate ion [87]. This is important to capture the right geometry of this planar molecule. The inversion potential used is defined as follows:

$$U(\phi_{ijkn}) = \frac{1}{2}k(\phi_{ijkn} - \phi_0)^2 \tag{2.14}$$

In Figure 2.4 we can see that ϕ_{ijkn} is the angle between the plane defined by three of the elements, in this case i, k and n, and the vector ij.

Intermolecular potential

The intermolecular potentials describe the interactions arising from the electrostatic and the Van der Waals contributions to the potential energy of the system.

Van der Waals Interactions

Short range interactions are important to capture the Pauli repulsion between atoms and the Van der Waals attractive forces induced by the formation of temporary dipoles. The former is essential in preventing the collapse of atoms on each other which would lead to an infinite electrostatic contribution to the potential energy.

A large number of different functionals can be used to define these interactions, however, during the following investigation the three types of potentials used include: the Lennard-Jones potential (LJ), the Buckingham potential (mainly for the mineral phase) and the n-m potential.

The LJ potential function has seen a widespread use for the description of intermolecular interactions. The reason lies in the simplicity of the potential and the successful replication of experimental data and first principle calculations for many different systems. It is usually expressed as:

$$U_{lj} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \tag{2.15}$$

where σ represent the distance at which the potential energy is equal to zero and ϵ is the depth of the minimum in the function.

The Buckingham potential is instead defined by an exponential function:

$$U_{Buck} = A \exp\left(\frac{-r_{ij}}{\rho}\right) - \frac{C}{r^6} \tag{2.16}$$

where A, ρ and C are parameters that depend on the interaction.

These two potentials are the only ones that will be used during the following analysis when investigating the energies, structures and dynamics of different configurations. An n-m potential was used to study the perfusion of Ar into a dehydrated ACC to investigate the potential presence of channels throughout the structure. The n-m potential is similar to a LJ potential but with a variable power law:

$$U_{nm} = \frac{E_0}{n-m}\left[\left(\frac{r_0}{r_{ij}}\right)^n - \left(\frac{r_0}{r_{ij}}\right)^m\right] \tag{2.17}$$

Here n and m are parameters which can be used, for instance, to modify the steepness of the Pauli repulsion wall.

Electrostatic interactions

The Coulomb potential between two charges is defined by:

$$U_{ele} = \frac{e_i e_j}{4\pi\epsilon_0 r_{ij}} \tag{2.18}$$

where ϵ_0 is the permittivity of empty space, r_{ij} is the distance between two atoms, and e_i and e_j are respectively the charges of atom i and j.

The Coulomb potential is a function of r_{ij}^{-1} and, compared to the Van der Waals forces discussed above, will affect elements at a much longer range. Thus truncation at short distance will lead to sizeable inaccuracies in the calculations but an explicit summation would be impractical as the sum between particles will need to be carried out between all atoms in the system and all their images. For periodic systems, alternative methods to calculate the electrostatic energy have been developed. One of these is the Ewald summation, which will be used in all the simulations described later.

For a cubic system, of length L and containing N atoms the Coulomb potential energy is obtained using:

$$U = \frac{1}{8\pi\epsilon_0} \sum_{n} \sum_{i=1}^{N} \sum_{j=1}^{N'} \frac{e_i e_j}{r_{ij} + nL} \tag{2.19}$$

where the sum runs over all periodic images n and over all particles, N, except $j = i$ when $n = 0$. Following this method, particle i will be interacting with all its periodic images. The equation is difficult to use in Molecular Dynamics simulations. The reason lies in its poor convergence as at longer distances the contribution from each atom decreases but the number of elements in the calculation increases.

In the method developed by Ewald, the infinite sum over all atom pairs is separated into two series, one calculated in real space, the other one in reciprocal space. Ewald noted that $\frac{1}{r}$ could be rewritten using:

$$\frac{1}{r} = \frac{\mathrm{erf(r)}}{r} + \frac{\mathrm{erfc(r)}}{r} \tag{2.20}$$

where the error function, erf, and its complement are given by:

$$\mathrm{erf(x)} = \frac{2}{\sqrt{\pi}} \int\limits_{0}^{\mathrm{x}} \exp -\mathrm{t}^2 \mathrm{dt} \tag{2.21}$$

$$\mathrm{erfc(x)} = \frac{2}{\sqrt{\pi}} \int\limits_{\mathrm{x}}^{\infty} \exp - \mathrm{t}^2 \mathrm{dt} \tag{2.22}$$

The first term in equation 2.20 is a slowly convergent sum at infinite distances, however it has a well-defined value at short r. In contrast the second term has an asymptotic behaviour as r goes to zero, but quickly converges to 0 as r goes to ∞. This term can be easily treated in real space where it quickly converges while the first term can be treated using a Fourier transform. In the method developed by Ewald each point charge is considered to be surrounded by a neutralising charge distribution as shown in Figure 2.5 *A* and *B* respectively. The charge distribution is represented by a Gaussian functional:

$$\rho(r)_{Gaussian} = \frac{q_i \alpha^3}{\pi^{\frac{3}{2}}} \exp\left(-\alpha^2 r^2\right) \tag{2.23}$$

where α determines the width of the Gaussian. The Gaussian distribution is subtracted from the particle point charge as shown in Figure 2.5 *b*:

$$\phi_{local} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{|n|=0}^{\prime} \frac{q_i q_j}{4\pi\epsilon_0} \frac{\mathrm{erfc}(\alpha|\mathrm{r_{ij}} + \mathrm{n}|)}{|r_{ij} + n|} \tag{2.24}$$

This sum will be easily computable in real space as it is short ranged. However, this will only give the unscreened fraction of the charges. We now need to add a compensatory charge distribution which will have the same form of the Gaussian distribution described earlier but will be of opposite sign. After a lengthy derivation, it can be shown that this is readily computable in Fourier, or reciprocal space using:

$$\phi_{long} = \frac{1}{2V} \sum_{k \neq 0} \frac{4\pi}{k^2} q_i q_j \exp\left(ik \cdot r_{ij}\right) \exp\left(\frac{-k^2}{4\alpha^2}\right) \tag{2.25}$$

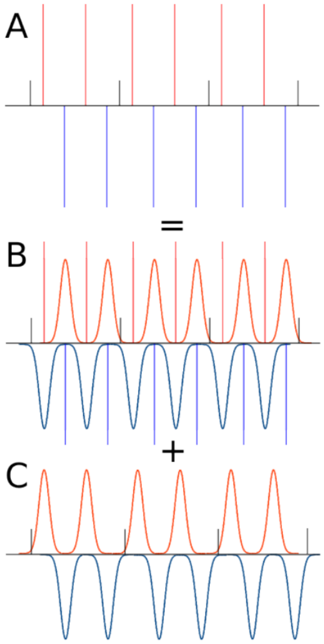


Figure 2.5: Ewald summation scheme. The original point charge description (a) is replaced by a sum of the point charges screened by a Gaussian distribution (b) and an additional Gaussian distribution (c) which cancels the effect of the first one.

where V is the volume of the system and k is the reciprocal space vector. An additional term arises due to the interaction between the Gaussian distribution and the point charge at the same site:

$$\phi_{self} = \frac{\alpha}{\sqrt{\pi}} \sum_{k=1} N \frac{q_k^2}{4\pi\epsilon_0} \tag{2.26}$$

The total contribution to the potential energy is then obtained by the sum of equations 2.26, 2.24 and 2.25.

2.3.2 Integration algorithm

The position of the atom at $t + \Delta t$, is obtained using the leapfrog version of the Verlet integration algorithm. It requires the position q^n, and force f^n of each element at the current timestep t while the velocities are obtained from t - $\frac{\Delta t}{2}$. The initial momenta of the particles are randomly assigned from a Maxwell distribution at the temperature of interest and thus all the parameters to initialize the simulation can be obtained. The following equations describe the leapfrog verlet algorithm:

$$v\left(t+\frac{\Delta t}{2}\right) = v\left(t-\frac{\Delta t}{2}\right) + \Delta t\frac{F(t)}{m} \tag{2.27}$$

$$r(t+\Delta t) = r(t) + v\left(t+\frac{\Delta t}{2}\right)\Delta t \tag{2.28}$$

where m is the mass of the atom and the force $F(t)$ is obtained via the potential energy of the system as described above. From the scheme presented above, it is clear that the velocity at the current timestep is not obtained directly. However, to calculate a number of properties in MD simulations both the velocities and the positions at all timesteps are needed. The velocity is computed using:

$$v(t) = \frac{v\left(t-\frac{\Delta t}{2}\right) + v\left(t+\frac{\Delta t}{2}\right)}{2} \tag{2.29}$$

2.3.3 Thermostat and barostat to run MD simulations in different ensembles

To maintain the pressure and temperature of the system a barostat and thermostat are usually used respectively. In the following work the Nosé-Hoover thermostat was selected.

The reason lies in the ability of this method to generate trajectories in both the canonical and isobaric ensemble. The other methods implemented in DL_POLY Classic [122], the Berendsen and Gaussian-constrains, produces properties that diverge from the canonical distributions.

The Nosé-Hoover thermostat works by extending Newton's equation of motion:

$$\frac{dr(t)}{dt} = v(t) \tag{2.30}$$

$$\frac{dv(t)}{dt} = \frac{f(t)}{m} - \chi(t)v(t) \tag{2.31}$$

$\chi(t)$ is the friction coefficient and is controlled by the following first order differential equation:

$$\frac{d\chi(t)}{dt} = \frac{N_f k_B}{Q}(T(t) - T_{ext}) \tag{2.32}$$

where $Q = N_f k_B T_{ext} \tau_T^2$ is the mass of the thermostat and τ_T is a time constant, or more precisely the relaxation time of the thermostat. Thus this value will determine the coupling between the heat bath and the system being simulated. N_f is the number of degree of freedom of the system.

Control over the pressure of the system has been achieved using a Nosé-Hoover barostat. This was implemented in DL_POLY classic using the Melchionna modification of the Hoover algorithm [123]. The equations of motions are now modified to couple a Nosé-Hoover thermostat and barostat.

$$\frac{dr(t)}{dt} = v(t) + \eta(r(t) - R_0) \tag{2.33}$$

$$\frac{dv(t)}{dt} = \frac{f(t)}{m} - [\chi(t) + \eta(t)]v(t) \tag{2.34}$$

$$\frac{d\chi(t)}{dt} = \frac{N_f k_B}{Q}(T(t) - T_{ext}) + \frac{1}{Q}(W\eta(t)^2 - k_B T_{ext}) \tag{2.35}$$

$$\frac{d\eta(t)}{dt} = \frac{3}{W}V(t)(\wp(t) - P_{ext}) - \chi(t)\eta(t) \tag{2.36}$$

$$\frac{dV(t)}{dt} = [3\eta(t)]V(t) \tag{2.37}$$

where $W = N_f k_B T_{ext} \tau_{\wp}^2$ is the mass of the barostat. χ is the barostat friction coefficient, R_0 is the system centre of mass, $\tau_{\wp}$ is the relaxation time of the pressure fluctuations, $\wp(t)$ is the instantaneous pressure of the system and V is the system volume.

For both the barostat and thermostat, the choice of the relaxation is an important factor. If this is too high, the coupling will be low and thus temperature or pressure will diverge from the selected value. In contrast, a too low value tend to lead to an incorrect distribution of velocities for a given temperature (non-maxwell).

2.3.4 Periodic image convention

Macromolecular systems contain $\approx 10^{23}$ elements, far more than can be simulated as a whole in MD simulations where we are constrained to a maximum of 10^9 elements. To circumvent the problem and be able to sample the properties of "bulk" systems, MD uses periodic boundary conditions so that no molecule is effectively in contact with a surface. Every time that a particle hits a boundary, it reappears on the other side of the box as demonstrated in Figure 2.6. Effectively the system is replicated infinitely around the central box.

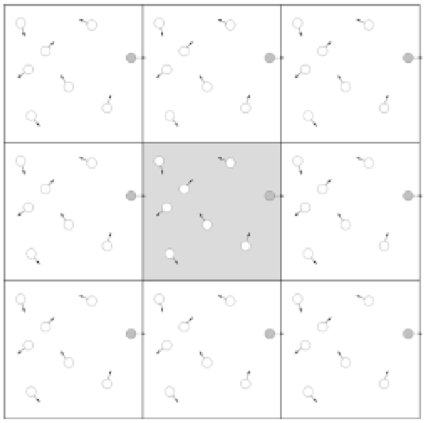


Figure 2.6: Periodic boundaries in Molecular Dynamics simulations.

2.3.5 Potential cut-off

The computationally most expensive part of running MD simulations is the calculation of the intermolecular forces. As specified above, the electrostatic and van der Waals terms depend on r^{-1} and r^{-6} respectively. The treatment of the electrostatic potential at long range has been described above. For the van der Waals forces the solution is to apply a taper to the potential function. Basically, instead of truncating the function at a precise point, which would lead to force and energy discontinuities across the simulation, we multiply the function by a term which depends on the distance between the particles. The following scheme was used:

$$U_{vdW} = f(r)U_{vdW} \tag{2.38}$$

$$\begin{cases} f(r) = 1 & \text{if } r < R_m \\ f(r) = f(r) & \text{if } R_m < r < R_c \\ f(r) = 0.0 & \text{if } R_c < r \end{cases} \tag{2.39}$$

Here R_m is the distance chosen to start the taper function and R_c is the distance at which the potential goes to zero. $f(r)$ is a function that smoothly changes from 1 to 0 between R_m and R_c. The first and second derivatives of this function need to be continuous to ensure that the integration algorithm works properly.

Ensemble averages of different properties can be obtained using MD simulations. However in most cases the simulations cannot sample all of phase space with the correct weighting during a computationally feasible timescale and thus an accurate average cannot be obtained. To circumvent this issue, a number of methods have been developed to sample configurational space in a more effective way. This will be discussed in the next section.

2.4 Rare events methods

Most rare events methods explore the partition function along one or more reaction coordinates. The enhanced sampling arise from either constraining the system to sample certain values of a reaction coordinate, or by impeding it to visit regions of the configurational space that have been already explored. Another set of methods, instead of sampling the partition function, sample the potential energy surface. These are based on first order kinetics and transition state theory [124].

Care needs to be taken in the definition of the reaction coordinate. It needs to be able to accurately differentiate the states of interest of the system under investigation and should not be computationally expensive. The choice of the reaction coordinate will be discussed in the next subsection. Then Umbrella Sampling will be briefly described as it is the method used in the following investigation.

2.4.1 Reaction coordinates

A large number of order parameters have been previously used to study phase transitions so this is by no means a comprehensive review of the literature. As mentioned before these need to be able to distinguish the different thermodynamic states of interest. However, a single reaction coordinate might not be sufficient. For example, the Q_6 Steinhardt bond order parameter has been used to study the crystallization of a Lennard-Jones sphere. This order parameter cannot distinguish the FCC from the BCC phase. Thus it will be able to drive the simulations without biasing polymorph selection, however, it will not be able to extract the factors that will determine this selection or discern between the two minima [125]. Santiso *et al.* recently published a research article discussing in details how to determine order parameter to analyse phase changes [126].

In a recent study, Raiteri *et al.* (2015) probed the free energy of ion binding as both a function of the distance between ions and their water coordination number [127]. Their result showed that for ions with a strong affinity to water, the one dimensional reaction coordinate including only the distance between ions led to a underestimation of the energetic barrier of binding [127]. Thus it is important to both accurately select a collective variable that can drive the system, but also to ensure that there are no important energy barriers on a free energy surface orthogonal to the one being sampled.

2.4.2 Umbrella Sampling

Umbrella Sampling was developed by Torrie and Valleau in 1977 to tackle problems where the sampling of a system was hindered by the potential energy surface [128]. The method developed allows us to recover the free energy surface by decreasing the degrees of freedom of the simulation by constraining the system to sample specific values along a reaction coordinates.

A system at temperature T and pressure P samples configurations from the isobaric isothermal ensemble:

$$f(q) \propto \exp\left(-\frac{U(q^N)}{k_BT}\right) \tag{2.40}$$

here q^N represents the atomic coordinates and k_B is Boltzmann constant. The vector q^N has a high number of dimensions, and thus it is much easier to analyse along a selected few degrees of freedom. The probability distribution for such a collective variable can be expressed as:

$$\rho(s) \propto \int \exp\left(-\frac{U(q^N)}{k_BT}\right)\delta(s - s(q^N)) \tag{2.41}$$

When this probability distribution is known the free energy landscape can be obtained by:

$$F(s) = -k_b T \log \rho(s) \tag{2.42}$$

Umbrella Sampling works by modifying the potential landscape by adding a potential that strictly depend upon the chosen collective variable. This potential is usually in the form of a restraint and in this study will be a harmonic potential of the form $V_{harm} = k * (s - s_0)$. So the potential is now $U(q^N) + V(s(q^N))$. This not only alows us to sample the regions of low probability, but saves computational time by only visiting regions of interest along the collective variable. Briefly, multiple simulations are set up along the order parameter s. The bias constrains the system to sample preferentially, proportionally to the bias and the potential landscape, values in the vicinity of s_0 in each simulation. This should sample a gaussian distribution around s_0 which overlap with its neighbours. We create a series of "umbrellas" along the reactions coordinate where we effectively sample biased distributions as shown in Figure 2.7. The biased probability of the distribution, $\rho'(s)$ is defined by:

$$\rho'(s) \propto \int dq^N \exp^{-\frac{U(q^N)+V(s(q^N))}{k_B T}} \delta(s - s(q^N)) \propto \exp^{\frac{-V(s(q^N))}{k_B T}} \rho(s) \tag{2.43}$$

and the resulting free energy biased landscape:

$$F'(s) = -k_b T \log \rho'(s) = F(s) + V(s) + C \tag{2.44}$$

The above equation shows that the additional potential is additive in the free energy. The constant of integration, C, is important when we want to recover the free energy landscape as it will shift the change in free energy obtained from the different windows

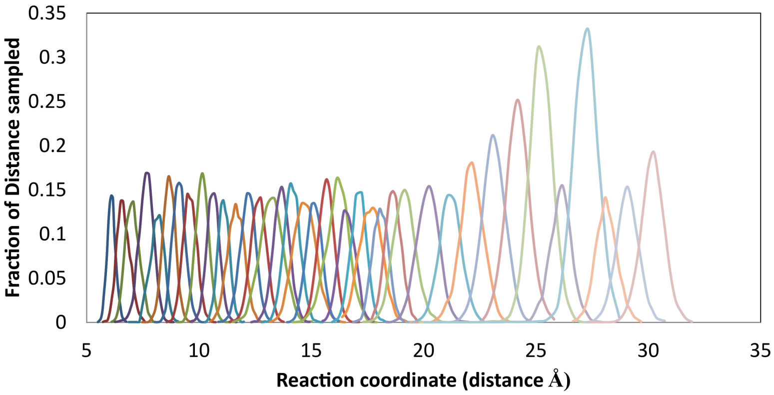


Figure 2.7: Distributions of s_0 values sampled along a reaction coordinate in multiple simulations. In this case the order parameter is the distance between two particles. Each distribution represent a different simulation with a specific s_0 value.

along the reaction coordinate. It is usually computed using a method called the weighted analysis histogram method (WHAM).

WHAM works by considering a series of simulations performed with restraints located at different positions on the reaction coordinate as shown in Figure 2.7. As mentioned above the constant of integration, C, is unknown. This depends on $V(s)$, the umbrella restraining potential and different simulations will have different offsets, C_i where i represents the different simulations along the order parameter s. The WHAM equations are then defined as:

$$\rho(s) = \frac{\sum_{i=1}^{N_{sims}} n_i(s)}{\sum_{i=1}^{N_{sims}} \exp\left(\frac{[C_i - V_i(s)]}{k_B T}\right)} \tag{2.45}$$

where C_i is determined selfconsistently using:

$$C_i = -k_B T \ln \left[\sum_{x_{bins}} \rho(s) \exp\left(\frac{-V_i(s)}{k_B T} \right) \right] \tag{2.46}$$

From these equations it is possible to recover the unbiased probability $\rho(s)$ by combining a series of restrained simulations that overlap along a reaction coordinate.

2.5 Analysis techniques

Some analysis methods will be used throughout the thesis. To avoid explaining them in each section, a brief description is here presented. Please refer to the references for more information.

2.5.1 Radial distribution function

The radial distribution function (RDF) describes the way by which atoms are radially structured around a reference atom within a system. It describes the probability of finding N atoms at a distance $r + \Delta r$ compared to the expected distribution for a random mixture [129]. Figure 2.8 shows the reference atom in brown and the atoms within $r + \Delta r$ are highlighted in blue. The RDF is defined by:

$$g(r) = \frac{V}{4\pi r^2 \delta r N} \sum_{i=1} \sum_{j=1, j\neq i} \delta(r - r_{ij}) \tag{2.47}$$

Where V is the volume and N the total number of particles within the system under consideration. The RDF can give insight into the structure. This property can be obtained

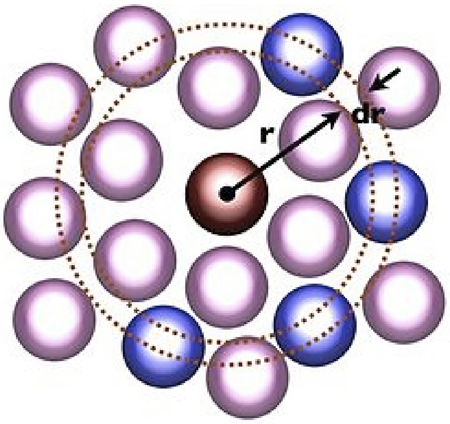


Figure 2.8: Definition of the radial distribution function (RDF). Here the reference atom is depicted in brown, the atoms within $r + \Delta r$ are highlighted in blue while the rest is depicted in light pink.

by X-ray experiment and so can be used to directly compare the outcome of simulation with experimental data.

2.5.2 Mean squared displacement

The mean squared displacement is a successful technique to obtain details of the movement of atoms within a system. This can be related to the diffusion of the atoms within the structure. Diffusion is often described in terms of Fick's law defined as [117]:

$$J = -D\frac{\delta c}{\delta r} \tag{2.48}$$

Where J is the flux ($\frac{mol}{m^2 s}$), c is the concentration ($\frac{mol}{vol}$) and r is the position. This relates the amount of atoms in a given volume and time to the concentration gradient within the solution. Fick's second law, further show how diffusion leads to changes in the concentration gradient with time:

$$\frac{\delta c}{\delta t} = D\frac{\delta^2 c}{\delta r^2} \tag{2.49}$$

This can be directly related to the average amount of motion that the molecules within a system undergo in a time $t + \Delta t$. In the case of the mean square displacement the absolute displacement is evaluated without considering the direction:

$$\lim_{t\to\infty} \frac{dr^2}{dt} = 6D \tag{2.50}$$

This information can be obtained in MD simulations by taking the position of particle at time t and calculating the square of the displacement between t and $t + \Delta t$.

Chapter 3

Simulation of calcium phosphate ions in solution: Analysis of nucleation pathways

3.1 Background

As mentioned in section 1.3.2, a number of recent studies have tried to probe the structures arising from the association of solutes prior to the nucleation event (i.e. the formation of critical size nuclei) [56, 76, 102, 104, 109]. In the case of biominerals, recent interest was spurred by the observation of structures considerably larger than an ion pair before nucleation in the calcium carbonate mineral system using analytical ultracentrifugation [3].

3.1.1 Early indications of complexes in solution

In the case of the calcium phosphate system the presence of complexes prior to nucleation has long been suspected. Posner was the first to suggest that the formation of both a calcium phosphate amorphous precursor and hydroxyapatite occurred through the aggregation of $Ca_9(PO_4)_6$ clusters later called Posner's clusters [130,131]. These have been subsequently investigated both experimentally and via computational methods [132–135]. The experiments of Onuma and Ito (using atomic force spectroscopy and dynamic light scattering to study the growth of hydroxyapatite) showed the presence of clusters in solution with sizes between 0.7 and 1.0 nm which they proposed to be $Ca_9(PO_4)_6$ clusters [134]. Additionally they observed that hydroxyapatite grew in steps of 0.8-1.6 nm suggesting that the clusters were the growth unit, supporting Posner's view [134]. Kanzaki *et al.* used *ab initio* methods to study the thermodynamic stability of various cluster sizes and aggregates of $Ca_3(PO_4)_2$, observing that a number of stable configurations existed on the potential energy surfaces of $[Ca_3(PO_4)_2]_3$ [132]. From the decrease of the energy as a function of the number of $Ca_3(PO_4)_2$ units they also proposed that the aggregation of these units could be an alternative to the formation of Posner's clusters. Their results, however, came from simulations carried in vacuum without water which could have important effects on the thermodynamics of the system. The water molecules usually stabilize the ion in solution by forming hydration spheres around it. These interactions arise from a favourable enthalpic interaction with the water molecule, even though the entropy is unfavourable due to the increase ordering of these water molecules. Additionally, the formation of neutral clusters as suggested in these studies does not explain the recent observation that the amorphous precursor to the nucleation and growth of calcium phosphate minerals is negatively charged due to a calcium deficiency [56, 57].

3.1.2 Precipitation of a calcium deficient amorphous structure

Recent *in vitro* experiments showed that, in contrast to previous observations where a calcium to phosphate ratio (Ca/P) between 1.1 and 1.5 [54] was obtained, ACP carries a negative charge due to a Ca/P ratio of 0.67 [56, 57]. A recent in vitro investigation of biomimetic calcium phosphate precipitation by Habraken *et al.*, which monitored the pH and composition of the solution, showed that the inorganic polymeric chains observed before nucleation using Cryo-TEM [109] were composed of monomeric units of calcium tri-hydrogen phosphate complexes, $[Ca(HPO_4)_3]^{4-}$, and calcium di-hydrogen phosphate, $[Ca(HPO_4)_2]^{2-}$ [56]. They suggested that the complexes interact through a double hydrogen bond within the polymeric chains. However, this could not be verified. Subsequent densification of the cluster would then occur by the addition of a further calcium ion leading to the precipitation of an amorphous structure composed of $[Ca_2(HPO_4)_3]^{2-}$ fractals [56]. The formation of complexes in solution was also observed using X-ray absorption near-edge spectroscopy (XANES) [136]. In this case the main species found was $[Ca(PO_4)_2]^{4-}$. Their results suggested that these did not participate in the nucleation event, which instead arose from the association of ion pairs.

A calcium deficient amorphous precursor structure appears to be important in vivo as well [59]. Analysis of osteoblast vesicles lining the growth front of murinae calvarial bone using electron dispersive spectroscopy showed a Ca/P ratio of 0.75 [59]. Interestingly, a recent investigation of the shark's tesserae (thin layer of mineralized tissue that surrounds the cartilage core of the skeleton in elasmobranchs) suggests that this amorphous precursor is composed of polyphosphate chains, which are later cleaved by alkaline phosphatase to produce the phosphate ions, further complicating the process [137].

3.1.3 Computational studies

Previous computational analysis of the formation of clusters for the calcium phosphate system using molecular dynamics simulations and *ab initio* methods has shown a wealth of clusters in solution. Zahn has looked at the thermodynamics of association of calcium and hydrogen phosphate ions using molecular mechanics and *ab initio* calculations [138]. He proposed that the stable initial complexes were composed of $(Ca_2HPO_4)^{2+}$, however other potential pathways were not examined [138]. De Leeuw *et al.* [139], and Ma *et al.* [140] looked at the interaction between calcium and phosphate ions in highly concentrated solutions in the presence of collagen using molecular dynamics simulations. They both observed the formation of calcium phosphate clusters. De Leeuw *et al.* analysed solutions with PO_4^{3-} anions, which are unlikely to exist in solution at a pH of 7.4 (the approximate pH of living systems). In contrast Ma *et al.* did investigate solutions containing HPO_4^{2-} and a number of clusters carrying a charge were reported [140]. However, Ma *et al.* obtained their force field for the HPO_4^{2-} molecule directly from the CHARMM database and the ability of that force field to model the bulk structures of any calcium phosphate minerals has not been investigated in depth [140]. Additionally we first need to understand the fundamental mechanisms of nucleation in calcium phosphates before we can consider the dramatic effect on both the structure and morphology related to the presence of organic molecules.

3.1.4 Aims and objectives

The aim of this study is to obtain a fundamental understanding of the process at a molecular scale. The formation of highly charged complexes in solution appears unlikely and thus a more thorough investigation is needed. In addition the different steps of the process cannot be directly observed using current experimental techniques and thus computer simulations could give further insights. The objective is thus to obtain information on the ion association process prior to the nucleation of calcium phosphate mineral phases. Here, we used molecular dynamics simulations to probe the formation, structure, stability and potential mechanisms of aggregation that can lead to the formation of a calcium deficient precipitate. Free energy profiles for both the formation and aggregation of the complexes will be presented. In addition we also sampled configurations with multiple complexes both by brute force molecular dynamics and using Umbrella Sampling. Lastly, to further investigate the potential differences between the simulations and experiments, an analysis of the effect of tris(hydroxymethyl)aminomethane (TRIS) have also been carried out. The results demonstrate the potential impact of this molecule on biomimetic experiments.

3.2 Methodology

3.2.1 Force field selection

The force field selected was obtained from Gale *et al.* [19]. It was developed to reproduce the thermodynamic properties of the ions in solution [19]. A similar method was previously used to obtain a force field for the calcium carbonate system and was shown to accurately capture the stability of the different polymorphs as well as their solubility in water [87,141]. In addition, there are no other force fields in the literature which incorporate the HPO_4^{2-} ion [142–146]. The force field parameters and an explanation of how they were obtained, are presented in Appendices A.2 and B respectively.

The water molecules were simulated using the SPC/Fw force field as suggested in the force field derivation [17]. The Na^+ and Cl^- counter ion interactions with the water molecules were obtained by fitting to their dissolution energy [19]. The interactions between Na^+ and the mineral phase were fitted using the disodium hydrogen phosphate crystal structure obtained by Catti *et al* [147]. Briefly, the energy of a number of structures with small displacements of an atom were calculated using density functional theory (DFT) within the Generalized Gradient Approximation (GGA) with the PBE functional using CASTEP [148]. GULP was then used to fit the interactions using the Newton-Raphson minimization method [149]. For the Cl^- no crystal structure is present thus the interactions of interest were obtained by fitting to a potential of mean force obtained by single point calculations in vacuum using again DFT within the GGA with the PBE functional [148].

3.2.2 Simulation method

All the simulations detailed in this study were performed using DL_POLY Classic version 1.9 [122]. The systems were initially equilibrated in an NVT ensemble at T=300 K, using a Nosé-Hoover thermostat with a relaxation time of 0.1 ps. The timestep was 1 fs and the simulations were run for at least 10^6 timesteps giving a total runtime of 1 ns. Subsequently the ensemble was changed to an NPT ensemble to allow the volume of the system to vary while keeping the temperature at 300 K. This was achieved using a Nosé-Hoover thermostat and barostat with relaxation times of 0.1 and 1.0 ps, respectively. For the production run, the timestep was again set at 1 fs. In the case of the Umbrella Sampling simulations, each window was run for at least 1 ns. The resulting distributions of the reaction coordinate showed a clear overlap along the order parameter. In the case of the standard Molecular Dynamics simulations, the production time lasted at least 4 ns. Calculations of the electrostatic interactions for all the simulations were treated using the standard Ewald sum.

Umbrella Sampling was implemented using the PLUMED plugin that interfaces with DL_POLY [150]. The integration along the reaction coordinate of the simulation was obtained by the Weighted Histogram Analysis Method (WHAM) using the software developed by Grossfield *et al* [151].

3.2.3 Formation of multiple calcium deficient complexes and their aggregation.

To study the formation of Ca^{2+} and HPO_4^{2-} complexes, the system geometry was prepared by randomly inserting one Ca^{2+} ion, three HPO_4^{2-} ions, four Na^+ ions and 9,000 water molecules into a simulation box (66.7 Å^3) using PACKMOL (tolerance set to 2.2 Å) [152]. The Ca-P distance was chosen as the reaction coordinate used to drive the Umbrella Sampling simulation. Starting with the ions well-separated, the distance was varied by 0.5 Å for each window until a value of 4.50 Å was reached and then by 0.25 Å down to 2.00 Å. After the first hydrogen phosphate reached the most stable free energy minimum, this contact ion pair was used as the starting configuration for docking the next phosphate ion to the complex. Before using Umbrella Sampling in this case, it was necessary to tether (tether force k=100 $kJ.\text{Å}^{-2}$) the Ca^{2+} ion to its current position until the new HPO_4^{2-} reached the desired distance. The tether on the Ca^{2+} ion was then removed and the Umbrella Sampling simulation started. This was necessary since, when the Umbrella Sampling simulation was started without a tether, the strength of the umbrella steering led to the dissociation of the Ca^{2+} and the first docked HPO_4^{2-}.

Umbrella Sampling was also used to explore the possible aggregation mechanisms of $[Ca(HPO_4)_2]^{2-}$ and $[Ca(HPO_4)_3]^{4-}$ complexes. Initially two complexes (either two $[Ca(HPO_4)_2]^{2-}$ or two $[Ca(HPO_4)_3]^{4-}$ complexes) were placed at a distance of 27 Å in a simulation cell containing 9,000 water molecules. The concentrations of the ions were slightly higher than those usually used experimentally to ensure that the simulations were feasible on a reasonable time scale. As the system carried a large negative charge, Na^+ ions were added to the system to make the overall charge of the simulation system neutral (these ions are always present during experiments examining the precipitation of calcium and phosphate minerals). In the case of the aggregation of two $[Ca(HPO_4)_3]^{4-}$ complexes, the system was simulated at three different ionic strengths (0.12, 0.17 and 0.30 M) to investigate how this could affect their aggregation. The ionic strength of the solution was increased by adding Na^+ and Cl^- ions to the system as required.

3.2.4 Molecular Dynamics simulation of clusters in solution

To further analyse the behaviour of the complexes described experimentally in solution, three $[Ca(HPO_4)_3]^{4-}$ units were placed in a simulation box containing 13,500 water molecules. The three units were placed at a distance of 12.5 Å from each other. This distance was chosen pragmatically to explore the possible presence of a metastable state at short to medium distances. The simulations contained eight Na^{+} and two additional additional Ca^{2+} ions to balance the charge of the $[Ca(HPO_4)_3]^{4-}$ units. The system was prepared in two different configurations; one in which the Ca^{2+} ions were close to the three $[Ca(HPO_4)_3]^{4-}$ units and another one where the Ca^{2+} ions were located at a minimum distance of 20 Å. The initial positions of the Na^{+} ions and the $[Ca(HPO_4)_3]^{4-}$ units were kept the same for both the simulations.

3.2.5 Densification of clusters

The nucleation of a secondary phase from solution occurs through a local increase in density of the solute ions. To gain an understanding of this process for calcium phosphate solutions, Umbrella Sampling was again employed. The radius of gyration was used as a collective variable. Five $[Ca(HPO_4)_3]^{4-}$ units were randomly inserted in a sphere with either five (ACP^{2-}) or seven (ACP^{1-}) additional Ca^{2+} ions. Additional Na^{+} ions were added to balance the negative charge. The system was then solvated using PACKMOL [152]. The Ca^{2+} ions in the simulations were selected as the ‘backbone’ of the molecules and were used for the calculation of the radius of gyration.

3.2.6 Investigation of complexes interactions with tris(hydroxymethyl) aminomethane (TRIS)

The structure for the TRIS molecule was obtained from the protein data bank. Its ionisation state was then set to that for a solution of pH 7.4 and the structure minimized using the GAMESS-UK code [153]. The minimized structure was used as input to obtain the partial charges of the atoms and the van der Waals interactions using Antechamber and the standard AMBER parameters [154]. Cross terms between the ions in the simulated solution and the TRIS molecule were obtained by using the method described by Freeman *et al* [21]. The system was prepared by inserting three $[Ca(HPO_4)_3]^{4-}$ complexes, two TRIS molecules, 20 Na^+ ions and 10 Cl^- ions in a simulation box. These were then solvated with 9,000 water molecules. The enthalpy of binding between the TRIS molecules and the $[Ca(HPO_4)_3]^{4-}$ complexes, $H_{TRIS-[Ca_3(HPO_4)_7]^{8-}}$ was obtained through the following scheme:

$$\begin{aligned} H_{2TRIS-complexes} &= H_{2TRIS-complexes-ions-water} - \\ (H_{complexes(-2(HPO_4)^{2-})-water} &+ H_{2TRIS-2(HPO_4)^{2-}-water} + H_{ions-water}) + 2H_{water} \end{aligned} \tag{3.1}$$

where $H_{2TRIS-complexes-ions-water}$ is the enthalpy obtained from the simulation of the whole system, $H_{complexes(-2(HPO_4)^{2-})-water}$ is obtained from the simulation of the complexes minus two HPO_4^{2-} ions (this is because these two HPO_4^{2-} ions were involved in the binding of two TRIS molecules), $H_{2TRIS-2(HPO_4)^{2-}-water}$ is obtained from the simulation of two TRIS molecules with the two additional HPO_4^{2-} ions that were bound to them and $H_{ions-water}$ include all the ions that were not present for charge balance in the other systems. None of the simulated systems were electrically charged.

3.3 Results and discussion

3.3.1 Formation of complexes.

Following the recent experimental results suggesting the formation of negatively charged $[Ca(HPO_4)_3]^{4-}$ complexes before nucleation and the subsequent formation of calcium deficient ACP, it is important to probe the mechanisms by which these complexes can form and their relative stabilities in solution [56]. The free energy profiles as a function of the distance between the Ca^{2+} and HPO_4^{2-} ions in the system are shown in Figure 3.1 for the formation of the following species; $Ca(HPO_4)$, $[Ca(HPO_4)_2]^{2-}$ and $[Ca(HPO_4)_3]^{4-}$. All the curves show a minimum at a Ca-P distance of 3.75 Å arising from the formation of the contact ion pair (CIP). This minimum is split into two configurations with a small barrier separating the two states. The inner configuration relates to the presence of a bidentate interaction (i.e. the HPO_4^{2-} ion is interacting with the Ca^{2+} ion via two $O_{phosphate}$) while the more stable interaction is due to a monodentate binding configuration. As HPO_4^{2-} ions are added to the Ca^{2+} ions the inner bidentate state is shifted to higher energies relative to the monodentate configuration. The minimum at 3.75 Å is separated from the solvent-shared ion pair (SIP) by a barrier at 4.73 Å (note that this barrier is negligible, i.e. below RT, for both the formation of the $Ca(HPO_4)$ and the $[Ca(HPO_4)_2]^{2-}$ complexes). For each additional HPO_4^{2-} ion, the height of this energetic barrier increases. This is to be expected as the Coulombic repulsion will increase due to the additional negative charges present. All the curves also show one additional shallow minimum at approximately 7.5 Å due to the formation of the solvent-separated ion pairs.

The free energy curve for the formation of the ion pair is highly exothermic with a value of -33.2 $kJ.mol^{-1}$. A similar behaviour was observed for the formation of $[Ca(HPO_4)_2]^{2-}$ complex with an energy change of -18.6 $kJ.mol^{-1}$. The further addition of a HPO_4^{2-} ion to this complex, leading to the $[Ca(HPO_4)_3]^{4-}$ complex described by Habraken *et al.*, also showed a minimum at 3.75 Å [56]. The effective energy difference between the two most stable states $[Ca(HPO_4)_3]^{4-}$ complex and solvent separated HPO_4^{2-} and $[Ca(HPO_4)_2]^{2-}$) is however, approximately zero and the barrier between the reactant and product state is 18.3 $kJ.mol^{-1}$, which corresponds to approximately 7.25 RT. The curve is also characterized by a long range tail which, levels off at approximately 27 Å as can be seen in Figure 3.2. Calculation of the Bjerrum length for divalent ions leads to a value of 27.9 Å. It is thus reasonable to suppose that the distance obtained by Umbrella Sampling corresponds to the Bjerrum length and can therefore be used as a reference point. This barrier will be difficult to overcome and only a small percentage, 4%, will be in the configuration proposed by Habraken *et al.* [56], while 96% (corresponding to the cumulative probability of finding the $[Ca(HPO_4)_2]^{2-}$ and HPO_4^{2-} at a distance between 27 Å and 5 Å) will be found as $[Ca(HPO_4)_2]^{2-} + HPO_4^{2-}$. The chemical equation with the relative free energy changes is shown in Figure 3.3. It is noteworthy that in an experimental study it was found that the solution was composed of 8% $[Ca(HPO_4)_3]^{4-}$ complexes while only 2% of the ions in the solution would be found as ion pairs indicating a much higher stability of the complexes compared to the ion pairs [56]. In addition a recent study, run under slightly different conditions, using synchrotron X-ray absorption near edge spectroscopy showed that a large percentage of the ions in solution were present as $[Ca(HPO_4)_2]^{2-}$ [136]. In our case we predict that all of the ions would be found as either $[Ca(HPO_4)_2]^{2-}$ or $[Ca(HPO_4)_3]^{4-}$ complexes with a higher percentage of the former. This lies within the range of charges observed by Habraken *et al.* for the complexes and agrees well with the recent study by Zhang *et al* [56, 136]. In addition, one has to note that in the paper of Habraken *et al.* ion pairs were taken into account in the calculations based on reported equilibrium constants [155]. In all reported literature, the values for the equilibrium constants of ion pairs were obtained taking into account that these are the only types of complexes present in solution under the conditions of the experiment. This assumption seemed to reasonably fit the mass balance calculations at the time, though their fits also depend on the choice of activity coefficient for the charged complexes, a value that is practically unknown. Such a restriction to one type of complex was also criticized by Niels Bjerrum who concluded that it was not resolved which type of calcium phosphate complex was present in solution [156]. Our simulations show that ion pairs are not important under the experimental conditions. Recalculating the data obtained by Habraken *et al.* [56] considering that $[Ca(HPO_4)_2]^{2-}$ ions are also present in the solution leads to a better fit to the $K_{eq}\gamma$ compared to previous calculations. The solution could contain up to 3.52% of $[Ca(HPO_4)_2]^{2-}$ and 4.42% $[Ca(HPO_4)_3]^{4-}$ type complexes in addition to 2% ion pairs.

As mentioned above, the formation of the initial ion pair is computed to be exothermic, with a free energy change of -33.2 $kJ.mol^{-1}$. There is a certain discrepancy between this value and the one previously obtained from experimental and computational studies [138, 155]. By using potentiometric measurements at 298 K, Chugtai *et al.* determined

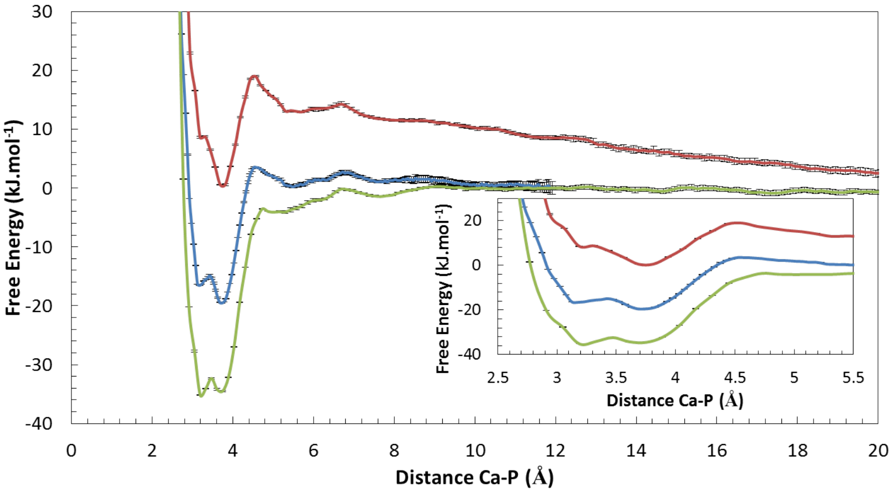


Figure 3.1: Free energy plotted as a function of the distance between Ca^{2+} and P (HPO_4^{2-} group). Here the zero of free energy is taken to lie at the furthest distance sampled which represents the separated ions. Green line represents $Ca^{2+} + HPO_4^{2-} \longrightarrow Ca(HPO_4)$, Blue line $Ca(HPO_4) + HPO_4^{2-} \longrightarrow [Ca(HPO_4)_2]^{2-}$, Red line $[Ca(HPO_4)_2]^{2-} + HPO_4^{2-} \longrightarrow [Ca(HPO_4)_3]^{4-}$.

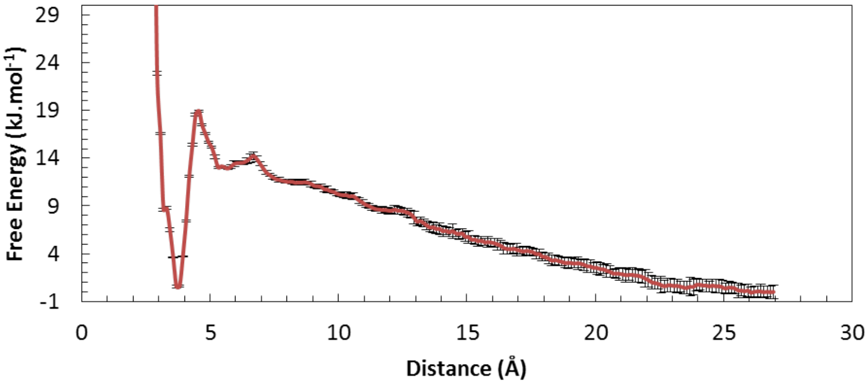
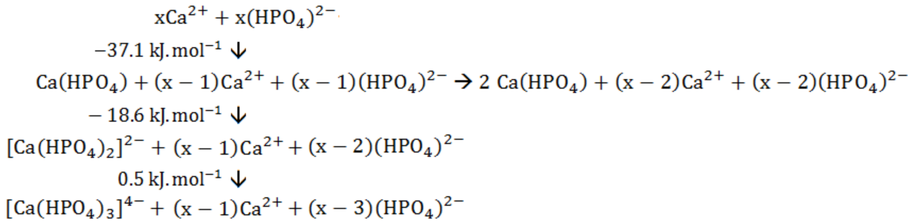


Figure 3.2: Change in free energy plotted against the distance between Ca^{2+} and P for the reaction $[Ca(HPO_4)_2]^{2-} + HPO_4^{2-} \longrightarrow [Ca(HPO_4)_3]^{4-}$.



$$xCa^{2+} + x(HPO_4)^{2-}$$

$$-37.1\ kJ.mol^{-1} \downarrow$$

$$Ca(HPO_4) + (x-1)Ca^{2+} + (x-1)(HPO_4)^{2-} \rightarrow 2\ Ca(HPO_4) + (x-2)Ca^{2+} + (x-2)(HPO_4)^{2-}$$

$$-18.6\ kJ.mol^{-1} \downarrow$$

$$[Ca(HPO_4)_2]^{2-} + (x-1)Ca^{2+} + (x-2)(HPO_4)^{2-}$$

$$0.5\ kJ.mol^{-1} \downarrow$$

$$[Ca(HPO_4)_3]^{4-} + (x-1)Ca^{2+} + (x-3)(HPO_4)^{2-}$$

Figure 3.3: Chemical equation with relative free energies changes.

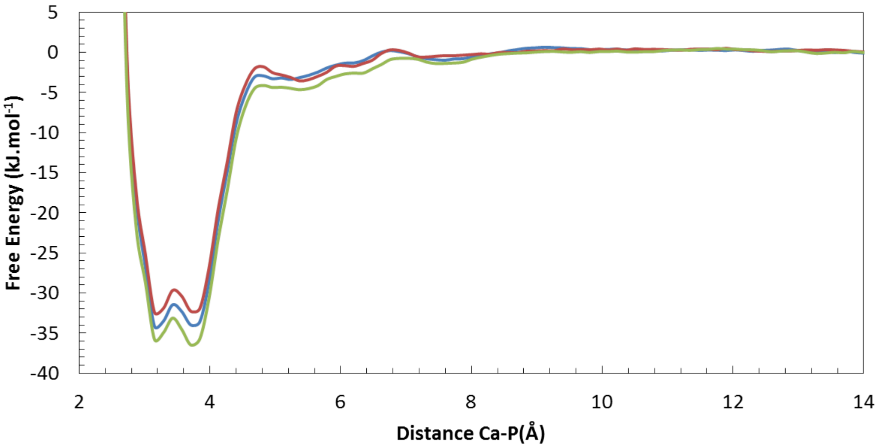


Figure 3.4: Change in free energy plotted against the distance between Ca^{2+} and P for the reaction $Ca^{2+} + HPO_4^{2-} \longrightarrow Ca(HPO_4)$. The green curve was obtained at T=310 K, the blue one at T=300 K and the red one at T=290 K.

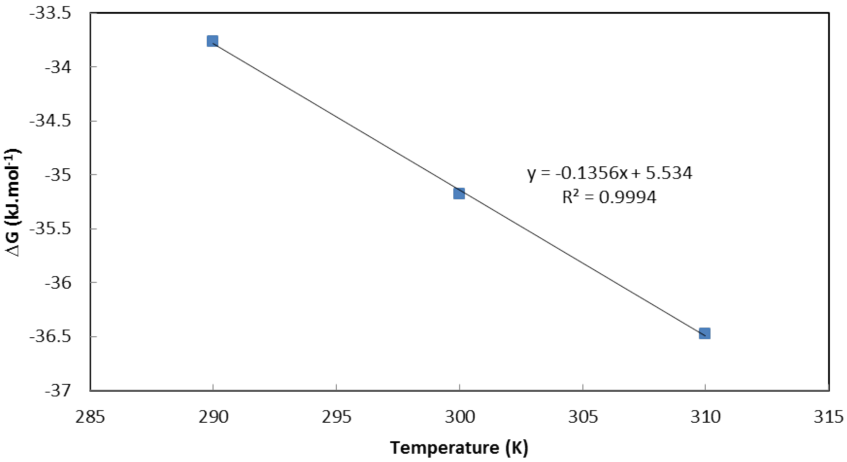


Figure 3.5: Change in free energy for the formation of the $Ca(HPO_4)$ ion pair plotted as a function of the temperature. The blue line is the line passing through the point, while the black one is a least square fit to the data points. R^2 describes the quality of the fit, with 1.0 being a perfect fit.

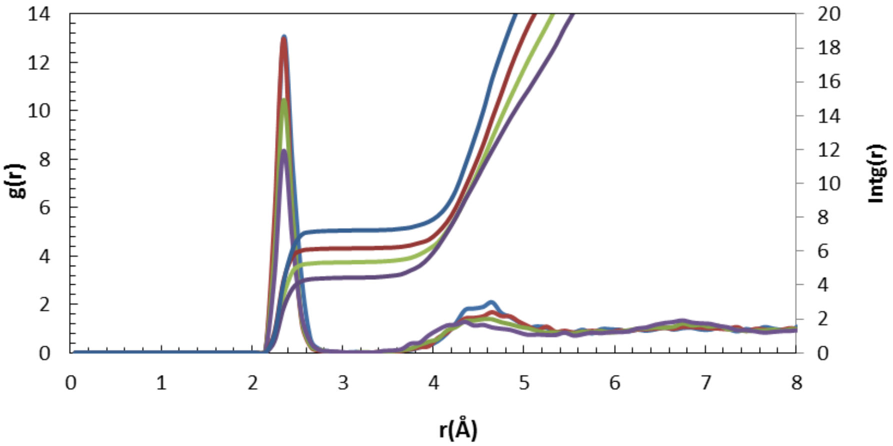


Figure 3.6: Variation of the Ca - O_{water} Radial Distribution Function (RDF) for complexes for each additional HPO_4^{2-} ion. The blue line shows the RDF for the Ca - O_{water}, the red line for the $Ca(HPO_4)$ unit, the green one for the $[Ca(HPO_4)_2]^{2-}$ and the purple one for $[Ca(HPO_4)_3]^{4-}$.

that the formation of the ion pair between Ca^{2+} ions and HPO_4^{2-} ions led to a change in the Gibbs free energy of -15.48 $kJ.mol^{-1}$ [155]. In contrast, *ab initio*/molecular mechanics simulations by Zahn gave a value of -27 $kJ.mol^{-1}$ [138]. Additionally X-ray near edge structure investigation has suggested that when a HPO_4^{2-} ion forms a contact ion pair with a Ca^{2+} ion, two O_{water} atoms are displaced from the coordination shell of the metal cation leading to a more centrosymmetric arrangement (6 O_{total} in the solvation shell compared to 7) [157]. However in our simulation, shown in Figure 3.6 and detailed in Table 3.1, only one O_{water} is displaced during the formation of the contact ion pair. Consequentially the Ca^{2+} will retain most of its enthalpy of hydration in addition to forming an additional 'bond' with an $O_{phosphate}$ which, in the current force field, has a higher partial charge (partial charge $O_{phosphate}$ = -0.89e vs O_{water} = -0.82e). Moreover, it is important to point out that in the calcium carbonate system, a decrease in the number of O in the coordination shell has rarely been observed during the precipitation of the amorphous precursor. A large number of experiments show that within hydrous amorphous calcium carbonate the Ca^{2+} will maintain a coordination shell of approximately 7.4 [14] suggesting that a similar behaviour in the phosphate system is to be expected. Calculation of the enthalpy and entropy change during ion pair formation can help shed some light on the root of the difference between the previous results and the current simulations. To obtain these values we ran the free energy calculations at different temperatures (i.e. 290, 300 and 310 K). The plot of the change in free energy as a function of temperature is presented in Figure 3.4. From this we extracted the change in free energy at a distance of 14 Å and plotted it as a function of temperature shown in Figure 3.5. The fit to this points permits us to obtain the change in enthalpy and entropy for the reaction following the standard equation:

$$\Delta G = \Delta H - T\Delta S \tag{3.2}$$

The values obtained from the multiple temperature simulations were 5.534 $kJ.mol^{-1}$ for the enthalpy and 0.136 $kJ.mol^{-1}.K^{-1}$ for the entropy. An unfavourable enthalpy change and a favourable entropy change is a general phenomenom for the formation of an ion pair as recently reviewed by Marcus and Hefter. The enthalpy of formation of a ion pair is unfavourable [155, 158]. Thus the formation of the ion pair is driven by a favourable entropy change, most probably arising from the release of water molecules from the coordination shell of the ions to the solution. Comparison of this results with the ones obtained by Calorimetric analysis by Chughtai *et al.*, '13.81 $kJ.mol^{-1}$ for the enthalpy and 0.096 $kJ.mol^{-1}.K^{-1}$ for the entropy' suggests that our model underestimates the unfavourable enthalpy contribution and overestimates the favourable entropy contribution compared to Chughtai *et al.* analysis [155]. However, as mentioned in the paragraph above, the association of ions into larger complexes was never taken into account in previous analysis and thus the experimental data of Chughtai et al. might not be a fully accurate set with which to compare the current results, and the values obtained by Zhan are much more similar to our present results [138, 155].

Further differences between the experiments and the simulations of Habraken *et al.* probably arise from not accounting for all the phosphate species that would be present in the solution at a pH value of 7.4. Under these conditions approximately half of the phosphate species would be present as $H_2PO_4^-$ (pKa=7.21). Another reason for the difference between the experiment and simulations could be the lower ionic strength used in the simulations

Table 3.1: Summary of water oxygens (O_{water}) in the hydration shell of each ion or ion pair. The change in coordination number is given relative to the species on the previous line of the Table.

Ion or calcium complex	Number of O_{water} in coordination shell	Change in O_{water} in coordination shell
Ca	7.22	-
$Ca(HPO_4)$	6.16	1.06
$[Ca(HPO_4)_2]^{2-}$	5.34	0.82
$[Ca(HPO_4)_3]^{4-}$	4.43	0.91

compared to the experiment, 0.06 mol/L compared to 0.2 mol/L, respectively. A higher ionic strength will screen the charges in solution, decreasing the Coulomb interactions and lead to a greater organisation of the water molecules. Additionally, the presence of other solutes in the experiments such as TRIS (used as a buffer) might affect the interactions. In a series of dynamic light scattering experiments, Onuma *et al.* looked at reference solutions to compare the size of the cluster observed when both Ca^{2+} and HPO_4^{2-} were present. The particle distribution obtained for the solution composed of NaCl, $K_2HPO_4.3H2O$, TRIS, and HCl showed a slightly larger particle size distribution when compared to the other reference solutions, potentially indicating that HPO_4^{2-} and TRIS interact to form small aggregates [134].

Table 3.2: Distance where change in free energy hange is greater than room temperature thermal energy for the aggregation of two $[Ca(HPO_4)_3]^{4-}$ as a function of ionic strength.

Ionic Strength (M)	Ca-Ca distance in Å
0.12	19.91
0.17	15.65
0.30	15.73

3.3.2 Aggregation of the different complexes in solution

From the results of the simulations detailed above, it is apparent that in solution there will be an equilibrium between two complexes, namely $[Ca(HPO_4)_2]^{2-}$ (found in higher concentration) and $[Ca(HPO_4)_3]^{4-}$ (suggested to be the monomeric unit of the polymeric chains from experimental observations) [56]. To form larger complexes, or aggregates, these species need to be able to interact in solution. Figure 3.7 shows the change in free energy for the interaction between two $[Ca(HPO_4)_3]^{4-}$ complexes for different ionic strengths as a function of the distance between the Ca^{2+} ions. All the curves presented show a similar behaviour, with an increase in free energy as the distance between the complexes is decreased. In the case of a solution with an ionic strength of 0.12 M there is no observation of a metastable state, however a slight inflection in the curve can be observed at approximately 7 Å. This is due to the direct sharing of one HPO_4^{2-} between the complexes. For an ionic strength of 0.12 M the change in free energy is greater than room temperature thermal energy (2.49 $kJ.mol^{-1}$) when the complexes are separated by a distance of 19.91 Å.

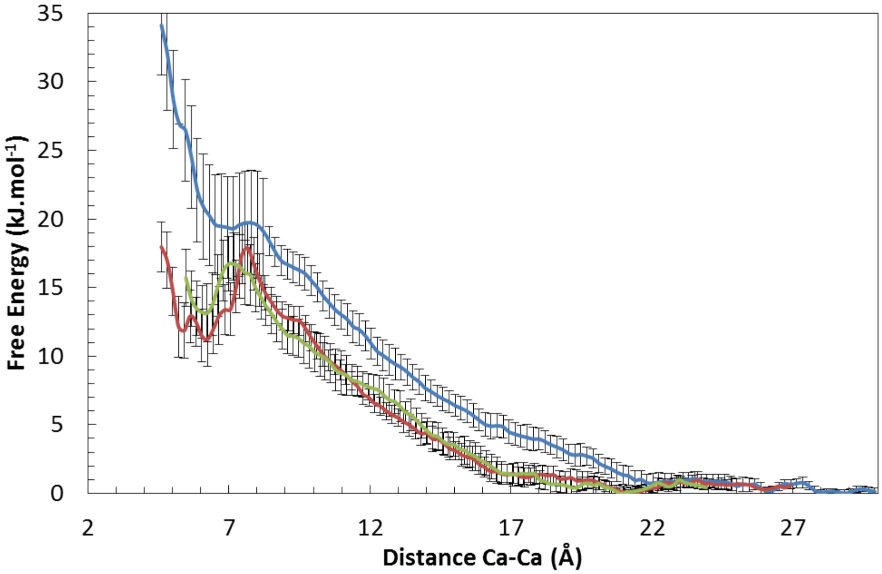


Figure 3.7: Free energy as a function of distance between the Ca^{2+} ions for two $[Ca(HPO_4)_3]^{4-}$ units. The blue curve was obtained from a solution with an ionic strength I = 0.12 M, the red for I = 0.17 M and the green for I = 0.30 M.

Increasing the ionic strength of the solution to 0.17 M (similar to the solution used in the experiments) slightly modifies the energetic landscape. The red and green curves in Figure 3.7 still show an increase in energy as the distance between the complexes is decreased but shifted to lower values. Interestingly, the inflection point observed in the curve obtained at an ionic strength of 0.12 M at a separation of 7Å between the complexes becomes a shallow metastable state separated by a barrier of approximately 4 $kJ.mol^{-1}$ (1.58 RT). This could potentially trap the complexes in this configuration for a small amount of time. Another interesting feature of increasing the ionic strength is the shift of the distance at which the room temperature thermal energy is overcome, allowing the complexes to come closer together (see Table 3.2). The complexes would, however, still be separated by four to five solvent shells (Figure 3.7). Further increasing the ionic strength to 0.30 M does not lead to further noticeable changes in the distribution. It is apparent that the addition of Na^{+} and Cl^{-} ions can indeed increase the charge screening in the simulation leading to a slightly less repulsive behaviour. However monovalent ions cannot further affect the aggregation of the complexes. The plateaux in the free energy curve is found at distances far too large for the complexes to form the double hydrogen bond interaction proposed by Habraken *et al.* Hydrogen bonding between the complexes was indeed observed in the simulations at approximately 12 Å [56] but was not observed to lead to the formation of a (meta)stable state.

$[Ca(HPO_4)_2]^{2-}$ complexes are the species that are most likely to be found in the solution simulated here and hence it is important to investigate their potential aggregation. In contrast to the $[Ca(HPO_4)_3]^{4-}$ complex, aggregation of two $[Ca(HPO_4)_2]^{2-}$ complexes

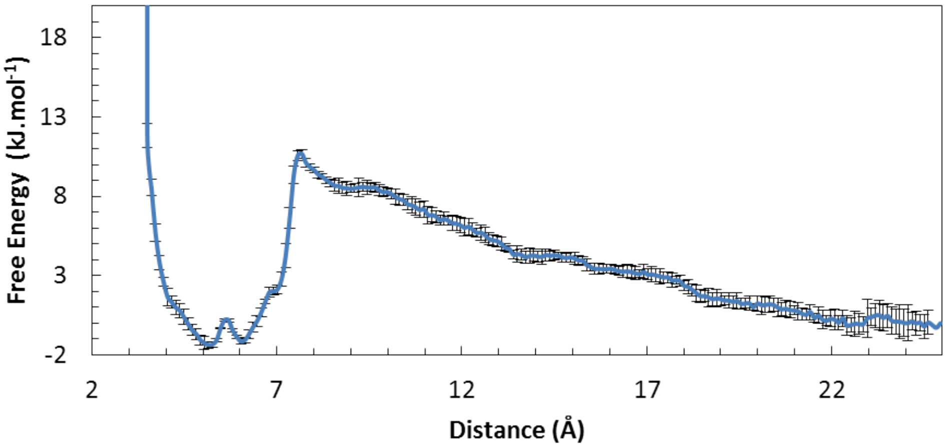


Figure 3.8: Free energy as a function of distance between the Ca^{2+} ions for two $[Ca(HPO_4)_2]^{2-}$ units.

showed a clear stable state between 5-7 Å (Figure 3.8). This minimum is split, with the different configurations representing the two stable states depicted in Figure 3.9. The stable configuration at shorter distances is characterised by direct phosphate sharing of two HPO_4^{2-} ions while the one at longer distance is due to the direct sharing of a single HPO_4^{2-} between the two complexes. The aggregate at shorter distances could be described as a $[Ca(HPO_4)_3]^{4-}$ complex bounded to an ion pair. It is also interesting to note that this aggregate will be only slightly more stable than the separate species (-1.2 $kJ.mol^{-1}$) with a long range barrier (18 Å) reaching a maximum of 10.6 $kJ.mol^{-1}$ separating the two states.

Following the nucleation analysis of Hu *et al.* the formation of states of comparable thermodynamic stability should not increase the barrier to nucleation and as such they

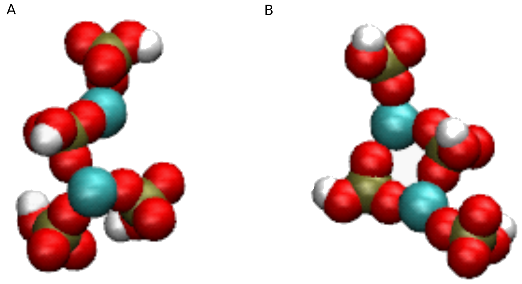


Figure 3.9: Examples of stable structures observed for the $[Ca(HPO_4)_2]^{2-}$ dimer. A) Shows the structure observed for the minima at 6 Å while B) shows the structures observed at 5 Å.

could potentially lie on the nucleation pathway. This is in contrast to the current view for the calcium carbonate system [3, 20, 102, 105]. In this case the association of ions into larger structures was observed to be driven by a constant decrease in the free energy both experimentally using potentiometric measurements and titration, and computationally using molecular dynamics simulations [3, 105].

Partially agreeing with previous results [56, 57], it appears that calcium and phosphate ions in solution can indeed form negative stable complexes, and that these will be able to either weakly interact or further aggregate. It is interesting to note that the association of two $[Ca(HPO_4)_2]^{2-}$ ions leads to an aggregate, $[Ca_2(HPO_4)_4]^{4-}$, which carries the same charge as $[Ca(HPO_4)_3]^{4-}$. This suggests that their interaction and the cross interactions between complexes might follow a similar free energy curve as obtained in Figure 5 for the aggregation of two $[Ca(HPO_4)_3]^{4-}$.

3.3.3 Molecular Dynamics simulations of clusters in solution

As the aggregation of two $[Ca(HPO_4)_3]^{4-}$ did not lead to any stable state at short distances, which was suggested experimentally, it is important to investigate how the additional presence of Ca^{2+} ions can affect their interactions and the potential for a metastable/stable state at short distances in their presence. In the case where the Ca^{2+} ions were absent or at a large distance from the three $[Ca(HPO_4)_3]^{4-}$ complexes, the complexes diffused away from each other until they reached an equilibrium distance which minimised the charge repulsion as shown in Figure 3.10 (a) and as expected from Figure 3.7. Additionally, the diffusion to larger separations led, in some cases, to the dissociation of an HPO_4^{2-} ion from the Ca^{2+}. In contrast, when the Ca^{2+} ions were in the proximity of the $[Ca(HPO_4)_3]^{4-}$ units, no loss of HPO_4^{2-} ions was observed (Figure 3.10 (b)). Interestingly, the additional Ca^{2+} ions interact with two $O_{phosphate}$ from two different HPO_4^{2-} in separate complexes leading to the formation of a negatively charged inorganic chain. The chain composition is $[Ca_4(HPO_4)_6]^{4-}$ which has the same Ca/P ratio as the ACP described by Habraken *et al.* [56] and Xie *et al.* [57] and only a slightly lower value than the one observed in osteoblast vesicles at the growth front of calvarial bone in murinae [59]. This result shows the molecular mechanism by which the complexes could potentially aggregate both in vitro and in vivo, and confirms experimental observations. As seen in the section above, the clusters could potentially be weakly interacting through five to six solvent shells. This weak interaction would lead to the local accumulation of negative charges that would attract Ca^{2+} ions due to the increasing Coulomb interaction. This could in turn potentially lead to the aggregation of the complexes as suggested experimentally. It is

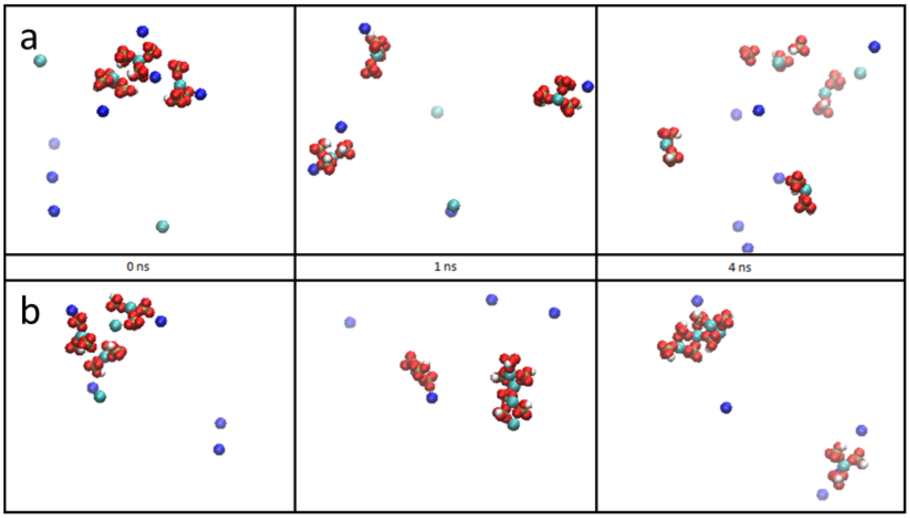


Figure 3.10: Snapshots of the simulation at different time steps. a) Shows the case where two Ca^{2+} ions (light blue spheres) were initially far from the complex. b) Two Ca^{2+} ions were initially in close proximity to the complexes. Ca^{2+} are depicted as turquoise spheres, P in green, O in red, H in white and the Na^{+} ions are shown as dark blue spheres. Water was removed for clarity from all the snapshots.

possible to speculate that the aggregation of two $[Ca_2(HPO_4)_4]^{4-}$ aggregates could follow a similar path as they carry the same charge. This would lead to a range of compositions very similar to those observed experimentally for amorphous calcium phosphate.

3.3.4 Densification of clusters

The result of the Umbrella Sampling simulations using the radius of gyration as a collective variable, showed a minimum in the potential of mean force for both ACP^{1-} and ACP^{2-}, as shown in Figure 3.11. The minimum is displaced towards a smaller radius of gyration for ACP^{1-} . This is sensible as the number of Ca^{2+} ions in the cluster is

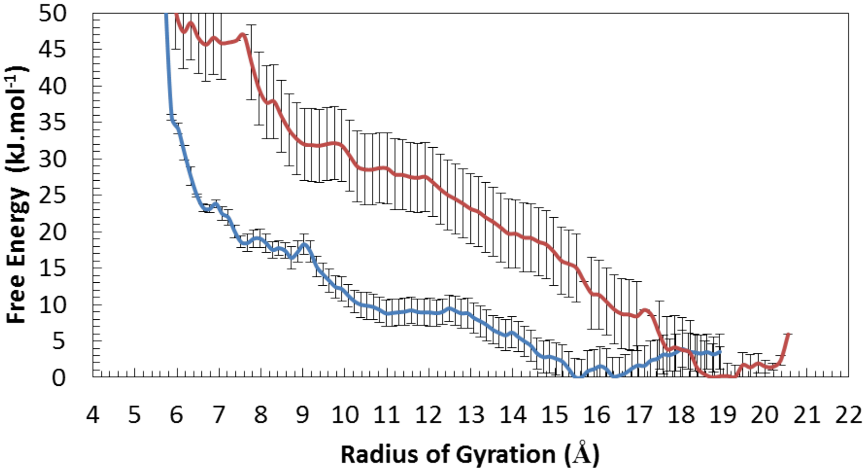


Figure 3.11: Free energy change as a function of the radius of gyration. The blue curve shows ACP^{1-} while the red one shows ACP^{2-}

increased (and consequentially the number of positive charges). It is interesting to note, however, that the curve is displaced only by approximately 2-3 Å. In both cases we see the formation of chain-like clusters and three additional units floating at some distance. In some cases these units complex with a Ca^{2+} ion. In the case where a larger number of negative charges are present in the cluster (ACP^{2-}), the potential of mean force increases much faster and does not show any stable state for denser structures. For ACP^{1-} the curve is flatter but no stable state is observed as the density is increased similar to the results obtained for the calcium carbonate system by Demichelis *et al.* [20].

3.3.5 Interaction of complexes with TRIS

The interaction of organic molecules with ions or mineral phases in solution is known to have an important impact both on the nucleation and the subsequent growth of the crystalline mineral or the amorphous precursor [40]. For example, polyaspartic acid, used as an *in vitro* model for non-collagenous proteins, has been shown to stabilize an amorphous calcium phosphate precursor in solution [42,113,159]. Similar effects have been observed for a number of non-collagenous proteins present in living systems like dentin phosphophoryn, a protein involved in the mineralization of bone and dentin [160]. Negatively charged proteins/residues have been mainly investigated *in vitro* due to their close association with biominerals. A recent paper by Cantaert *et al.*, however, showed that the positively charged poly(allylamine hydrochloride) could also have a dramatic effect on the nucleation and growth of calcium carbonate [113]. TRIS is a positively charged organic molecule that has been widely used as a buffer agent during *in vitro* biomimetic investigations of the nucleation and growth of calcium phosphate minerals [56,134,161]. In a recent investigation it was observed that TRIS delayed the nucleation of calcium phosphate minerals [56], thus it is important to understand how it could potentially interact with the clusters/ions and affect their stability.

The results of the molecular dynamics simulations showed that the TRIS molecules were able to act as a substrate for the localized accumulation of Ca^{2+} and HPO_4^{2-} ions. The two TRIS molecules in the solution aggregated via an interaction with two HPO_4^{2-} ions which bridged the two molecules by forming hydrogen bonds with both the hydroxyl group and the amine group (Figure 3.12). The Ca^{2+} and HPO_4^{2-} ions in the complexes interacted with the TRIS molecules through the formation of hydrogen bonds between one of the HPO_4^{2-} and the amine group of the TRIS molecule. This composite (organic+inorganic) is stable throughout the remainder of the simulation (4 ns). The binding enthalpy of the calcium phosphate complexes to the TRIS molecule was found to be -284.00 $\pm$86.14 $kJ.mol^{-1}$. It is important to note, as mentioned in the methods section, that Na^+ and Cl^- ions are also present in the system and do interact with the complexes and the TRIS molecules and thus play a role in the observed interaction. However, as in the other sections, the Na^+ and Cl^- ions are never truly bound to the complexes or TRIS, but sporadically interact via either direct interactions or through the water solvent shell.

These results highlight multiple potential roles of the TRIS molecule in solution. First of all, TRIS and HPO_4^{2-} ions are able to form small aggregates which appear to be highly stable and might be the cause of the slightly larger particle sizes observed by Onuma *et al.* using dynamic light scattering [134]. Additionally, this aggregate appears to be able to promote the stabilization of a negatively charged complex composed of 3 Ca^{2+}, 9 HPO_4^{2-} and two TRIS molecules. During the Umbrella Sampling investigation of the aggregation of two $[Ca(HPO_4)_3]^{4-}$ complexes, the increase of monovalent ions in the simulation did not induce the formation of any stable structure. Therefore the TRIS molecule is able to lead to a densification of the $[Ca(HPO_4)_3]^{4-}$ cluster through interaction with multiple functional groups and the formation of an interface, which decreases the interfacial energy of the negative complexes.

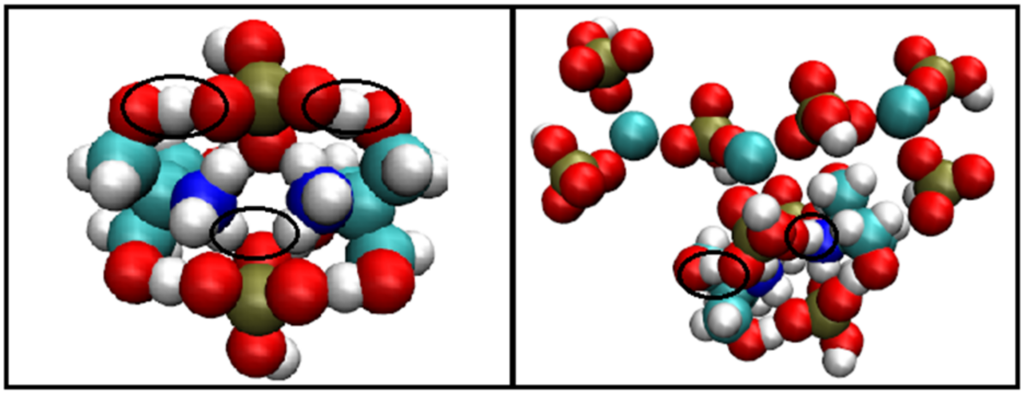


Figure 3.12: a) TRIS molecule bridged by two HPO_4^{2-} ions. b) Interaction of TRIS molecules with $[Ca(HPO_4)_3]^{4-}$ complexes. Water, Na^+, and Cl^- were removed from the image for clarity. A number of Na^+ ions are also observed close to the complexes in the simulations.

3.4 Nucleation pathways

From the results detailed above, there are a number of potential conclusions on the nucleation mechanism that can be drawn. Before we start, it is important to point out that we are not claiming to have defined the definite pathway to nucleation and crystallization. In this chapter the aim was to investigate, at a molecular level, the recent experimental results reported for the nucleation and growth of calcium phosphate minerals by Habraken *et al.* and Xie *et al.* and obtain a fundamental understanding of the mechanism involved and the potential molecular species which could play an important role [56, 57]. The observation of a calcium deficient amorphous precursor is surprising as it is unusual to think of natural processes evolving through electrically unbalanced reactions. The calcium phosphate complexes, as defined by Habraken *et al.* are here shown to be stable and potentially long lived in solution. If we consider the simulations carried at higher ionic strength it is plausible to speculate that in such solution and with the addition of H_2PO^{2-} to account for the full speciation, the stability between the $[Ca(HPO_4)_3]^{4-}$ and $[Ca(HPO_4)_2]^{2-}$ complexes could be reversed to what was found in our simulations, where the $[Ca(HPO_4)_2]^{2-}$, in contrast to the $[Ca(HPO_4)_3]^{4-}$, will be the major species in solution. The stability order our simulations predict was recently observed using X-ray absorption spectroscopy by Zhang *et al.*, using a solution with a different supersaturation [136]. However, the complexes present in the solution will be dependent upon a number of factors that will include the solution composition, the presence of organic molecules (i.e. TRIS) and the physical constrains applied to the solution.

This study shows, in contrast to the calcium carbonate system, that for the formation of complexes and their aggregation, energetic barriers are present along the free energy landscape, and as such the phosphate complexes and their aggregates in solution will tend to be long lived and could form a more defined interface with the solution. The work by Hu *et al.*, later used by Habraken *et al.* to explain how the formation of the clusters would affect the barrier to the nucleation of ACP, used a modified version of Classical Nucleation theory (CNT) [56, 107]. Hu *et al.* argued that if, in contrast to CNT, we do not assume that the interfacial energy is the same as the bulk phase, but instead changes as a function of the size of the forming nuclei so that it reaches a negligible value for the ion pair. Then the free energy barrier for the transition will be scaled by:

$$\Delta G_{ex} \propto \frac{\alpha^3}{(\sigma \pm C)^2} \tag{3.3}$$

Where C is a factor that depends on various parameters; including the shape of the cluster and the cluster radius. Most importantly the excess free energy of the cluster. α is the interfacial free energy and σ measures the favourable energy arising from the association of ions in solution. The plus and minus sign in the denominator depends on whether the energy minimum occupied by the cluster is of higher or lower energy compared to the individual ions or ion pairs in solution. In the latter case a minus sign will apply and the barrier to nucleation will be increased. In the former case the barrier will be decreased and the complexes is likely to lie on the nucleation pathway.

Figure 3.13 depicts the change in free energy as a function of the hydrogen-phosphates binding to the cluster. In Figure 3.13 the free energy axis, the free energy barriers and the overall energy differences are not to scale but are here shown as a guide to the eye. A number of clusters are seen to be stable even if separated by small barriers, as presented in Figure 3.13, with generally a ratio of $Ca^{2+} : 2(HPO_4)^{2-}$. The TRIS molecule is seen to interact and stabilise the clusters as shown in Figure 3.13. It should be stressed, however, that a question mark should be added as we do not know precisely the change in free energy. The same stabilising influence is not seen for monovalent ions which carry the same electric charge. The effect of the TRIS on the water structure and its interaction via the hydroxyl and amine functional group with the $O_{phosphate}$ appears to be essential in the stabilization of a cluster composed of 3 Ca^{2+} and 9 HPO_4^{2-}. These simulations give further support to the recent results of Xie *et al.* and Habraken *et al.* by showing the molecular mechanisms behind their observations and highlighting further potential pathways [56, 57].

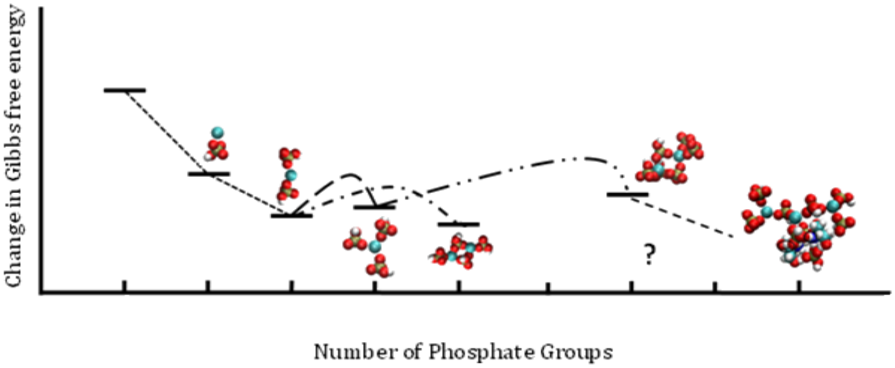


Figure 3.13: Representation of the change in free energy change as a function of number of phosphate in the complex. The axis describing the change in free energy is not to scale and should be used as a guide to the eye. The phosphorous is here depicted in green, the oxygens in red, the hydrogens in white, the calciums in blue, the carbons in grey and the nitrogens are depicted in dark blue.

3.5 Conclusion

The results obtained here from Molecular Dynamics simulations provide further evidence that negatively charged complexes of calcium phosphate can indeed form and be stable in solution confirming previous experimental observations. Our simulations suggest a dynamical equilibrium between the $[Ca(HPO_4)_3]^{4-}$ and $[Ca(HPO_4)_2]^{2-} + HPO_4^{2-}$ states, which fits within the range of charges observed experimentally. From the free energy curves, we calculated the population of the different clusters showing that 4% of the ions in the solution will be present as $[Ca(HPO_4)_3]^{4-}$ while 96% of the population will be found as $[Ca(HPO_4)_2]^{2-} + HPO_4^{2-}$. The higher amounts of $[Ca(HPO_4)_2]^{2-}$ found in our simulations compared to experiments could arises from the lower ionic strength of the simulations, 0.06 compared to 0.2 used experimentally. The higher ionic strength was not investigated as the current set of simulations already showed a minima in the free energy function and that the calculations require substantial amount of time to be performed.

When multiple $[Ca(HPO_4)_3]^{4-}$ where in proximity, the additional presence of Ca^{2+} ions led to the formation of chains, which were stable for extended period of time. These inorganic chains had the same composition found for amorphous calcium phosphate experimentally in two separate studies [56, 57]. Additionally, vesicles in osteoblasts present at the growth front of bone showed a Ca/P ratio of 0.75 suggesting the potential importance of this pathway for the *in vivo* mineralization of calcium phosphate.

The aggregation of the complexes leads to configurations that are either on the same level of stability or higher in energy suggesting that these could potentially lie on the nucleation pathway. This would potentially decrease the barrier to nucleation as originally suggested by Hu *et al.* and subsequently used by Habraken *et al.* to explain how amorphous calcium phosphate could form at their studied supersaturation [56, 107]. The formation of complexes, and taking into account the interfacial energy that these carry was shown to decrease the nucleation barrier to a reasonable value for amorphous calcium phosphate to precipitate.

As in this study we observed a much higher population of $[Ca(HPO_4)_2]^{2-}$ complexes and increasing the ionic strength was observed to have only a slight effect on the free energy curve, it raises the question of the potential role of TRIS. This molecule was able to stabilize the localized association of three $[Ca(HPO_4)_3]^{4-}$ complexes. The TRIS molecule carries a positive 1e charge, which is the same as Na^+. However, the presence of various functional groups (amines and hydroxyls), which were observed to interact strongly with the hydrogen phosphate ions, was able to stabilize a large negative charge for extended periods of time. This is one of the first time that the effect of this buffer on the precipitating phase has been analysed in detail showing that it is not just a passive observer but actively participates in the precipitation of calcium phosphate minerals.

3.6 Future work

A few months before the end of my studies, our collaborators at Curtin University detected an issue with the water structure around the hydrogen phosphate anion. As shown in Figure B.1 in Appendix B the water density in the two models is considerably different. In the current model the water hydrogens lie in between the oxygens of the hydrogen phosphate anion leading to an unrealistic hydrogen bond angle. They form an hydrogen bond with two oxygens contemporarily which is unlikely and distinct to *ab initio* calculations as explained in more details in Appendix B. The reason for the erroneous water structure in the current potential was due to the intermolecular interaction between the hydrogen of the water molecules and the oxygens of the hydrogen phosphate anions. These have been removed in the new model.

The new force field, however, predicts a dissolution energy 200 $kJ.mol^{-1}$ more favourable than the best estimate, -1140 $kJ.mol^{-1}$, obtained by Gale *et al* (see Appendix B) [19].

The one obtained with the previous potential was 15 $kJ.mol^{-1}$ more favourable compared to the best current estimate. As the two force fields demonstrate a substantially different interaction with water, comparing the results presented above against one obtained with the new model could lead to a better understanding of how the system should actually behave.

The change in free energy for the formation of the ion pair using the new force field is considerably lower compared to the old model. Now the value is closer to experimental result obtained by Chughtai *et al* [155], -15 $kJ.mol^{-1}$ [19]. This force field also predicts, as the old one, a favourable change in Gibbs Free energy for the formation of the $[Ca(HPO_4)_2]^{2-}$ complex. The values obtained were -6 $kJ.mol^{-1}$ for the new force field [19] and -18 $kJ.mol^{-1}$ for the old model. This provides strong evidence that this ionic complex is important in calcium phosphate solutions. They could either be involved in the subsequent precipitation or decrease the ionic concentration by forming a stable unit. In EDXA experiments, these complexes were suggested to be present in solution but not to take part in the nucleation process, which they suggested occured by the aggregation of ion pairs and ions [136]. In both cases, it suggests that the equilibrium constant for the formation of an ion pair should be recalculated considering the potential presence of these complexes.

Formation of the $[Ca(HPO_4)_3]^{4-}$ complex could not be studied with the new potential. Already in the case of the $[Ca(HPO_4)_2]^{2-}$ complexes attachment and detachment events could be observed during simulations. This results from the low energy difference, approximately twice as large as kT at room temperature, and the absence of a transition energy

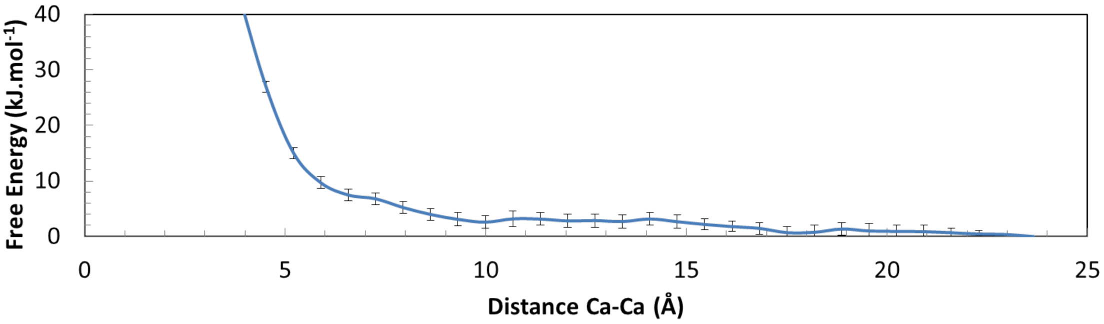


Figure 3.14: Free energy as a function of distance between the Ca^{2+} ions for two $[Ca(HPO_4)_2]^{2-}$ units.

barrier.

Aggregation of two $[Ca(HPO_4)_2]^{2-}$ complexes was also studied. As shown in Figure 3.14 the free energy change as a function of the Ca^{2+} ions distance is below RT thermal energy fluctuations up to a distance of 9.5 Å. At the ionic strength analysed in the current simulations the shallow metastable states are not separated by a barrier large enough to suggest they could be trapped in an aggregate state. Simulations at different ionic strength should be analysed to verify the effect on the free energy change. It would also be interesting to see how the presence of additional complexes will affect the interaction. Observations at different windows of the Umbrella Sampling simulation using the new force field showed that the hydrogen phosphate anions often exchange between coordination shells, and can also swap between calcium ions when the distance is short. Detachments events were also observed and were more frequent than with the old force field. The cause is the lower free energy minimum, and thus lower barrier to detachment.

TRIS interaction with the complexes using the new force field led to a slightly unfavourable interaction energy, 60.67 ± 79.34 $kJ.mol^{-1}$. The standard error, however, is large enough to reach favourable values. Additional simulation time is needed to obtain a more accurate estimate.

In future, it would be interesting to understand how the complexes further aggregate and the evolution of the structure and chemistry with complex size. Long time-scale simulations of ions in solution at a pH of 7.4 could give speciation values for the different complexes in solution. Additional complexes not sampled during this investigation might be present and could be more important.

If such a benchmark model could be obtained and matched to experiment, it would be possible to study the effect of organic additives at different steps during the process. The results obtained above with the addition of the TRIS molecule clearly show that organic additives could have an unexpected impact on the intermolecular interactions between ions. The results could lead to new design principles to control mineral growth.

Chapter 4

Structure and properties of hydrous amorphous calcium carbonate

4.1 Background

4.1.1 Introduction

As mentioned in Section 1.2.5, the formation of biominerals is often preceded by the precipitation of metastable precursors. The process has now been observed in both vertebrates and invertebrates [53, 88, 162]. Hydrated amorphous calcium carbonate (ACC) is an example of such a phase and is the precursor to calcite [15, 89], aragonite [10, 90] and vaterite [10] in many living organisms. Control over structures, polymorph selection and properties is achieved by a complex interplay between the solution chemistry [86], the presence of inorganic [11, 101] or organic [163–165] molecules and the physical conditions [166, 167] under which the phase precipitates. These processes have inspired new synthetic methods including the replication of the sea urchin skeletal plate using solid substrates [168] or the fabrication of micro-lens arrays by self-assembly from amorphous calcium carbonate (ACC) mediated by organic molecules [169]. However, before being able to use such approaches effectively, the mechanisms of morphological control and polymorph selection must be fully understood. Since ACC is a common starting point from which polymorph selection occurs in carbonate biominerals, a deeper knowledge of the structure and properties of ACC is a key step in gaining this understanding.

4.1.2 Experimental studies of ACC

Experimental evidence, from analysis of either biologically or synthetically prepared samples, indicates that ACC can exists in a number of different forms. These were initially distinguished by their stability, which is dependent upon the presence or absence of water within the structure. Although less hydrated ACC has been reported in some systems [12], there is now a consensus that the biologically relevant composition is approximately $CaCO_3.H_2O$. ACC with this composition has been observed to persist for days to months, while the anhydrous counterpart readily transforms into one of the final crystalline polymorphs [11, 170].

Previous analysis by X-ray total scattering showed that both short and medium range order are present within ACC. No correlation, however, in the pair distribution function could be found for distances greater than 10 Å, suggesting the absence of long range order [13, 84]. Ca K-edge extended X-ray absorption fine spectroscopy (EXAFS) has been used to determine the first coordination shell of the calcium ion. Approximately seven oxygen atoms were found in the calcium coordination shell on average, with typical bond lengths between 2.40 and 2.50 Å but the details proved to be highly dependent on the sample [4, 84].

Proton solid-state nuclear magnetic resonance showed that two different types of water are present within the structure. One is largely immobile and labelled as structural, whereas the other showed an increased mobility [84, 85, 171]. The location of this water appears, again, to be dependent on the sample origin. In synthetic samples the mobile water is in contact with the carbonate ions [84] while in biogenic samples no mobile water is identified close to the carbonate ions. Instead it was hypothesized by Reeder *et al.* to be segregated in channels [85].

As already discussed in Section 1.2.5, ACC can already have features resembling the final polymorph. This structural versatility was called polyamorphism and was initially suggested by Addadi et al. [12] for ACC. The suggestion arose from the observation of different short range-orders in living organisms such as in the shell of the snail Biomphalaria Glabrata [10]. The adult shell is mainly composed of aragonite, however, during the growth stage ACC is the first mineral phase that precipitates. K-edge EXAFS showed aragonitic order in the ACC present in the shell [10]. In another study, Politi *et al.* analysed the ACC present in the spicules of sea urchin embryos which in later stages of growth transforms into calcite. Infrared spectra confirmed that the mineral phase was amorphous and in complement with X-ray absorption spectroscopy (XAS) showed that the order present was reminiscent of calcite [15, 172].

4.1.3 Modelling studies of ACC

Molecular modelling has also been used to study both anhydrous [87, 173, 174] and hydrated ACC [4, 87, 105, 175–177].

Quigley *et al.* used metadynamics to sample the free energy landscape of calcium carbonate nanoparticles of various sizes using the potential model developed by Pavese *et al.* [173,174, 178]. When the system was simulated at constant volume, they found minima pertaining to ACC and vaterite in the free energy landscape. However, when the pressure of the system was kept constant, a minimum for calcite was found for nanoparticles with a size as little as 2 nm [173,174]. The observation that constant volume conditions favoured ACC formation was in good agreement with experimental observations [167]. By contrast, Raiteri and Gale, using a new force field fitted to thermodynamic parameters, showed that nanoparticles of ACC were more stable than calcite for diameters up to 3.8 nm [87]. Additionally, they found that as the ACC particle grew the amount of water in the structure increased [87]. ACC could thus have a heterogeneous structure. Bushuev, Finney and Rodger recently sampled the stability of bulk ACC as a function of water content and found a minimum in the potential energy between one and six water molecules per calcium carbonate formula units for this phase, agreeing with previous work [179].

The structure of ACC has also been studied using molecular modelling. Goodwin *et al.* produced a model of hydrous ACC using Reverse Monte-Carlo [4]. They fitted a system composed of $CaCO_3.H_2O$ formula units to an experimentally obtained ACC structural factor. Their analysis showed the presence of percolating water channels rich in carbonate ions while the bulk structure was richer in Ca^{2+} ions. A subsequent study by Singer *et al.*

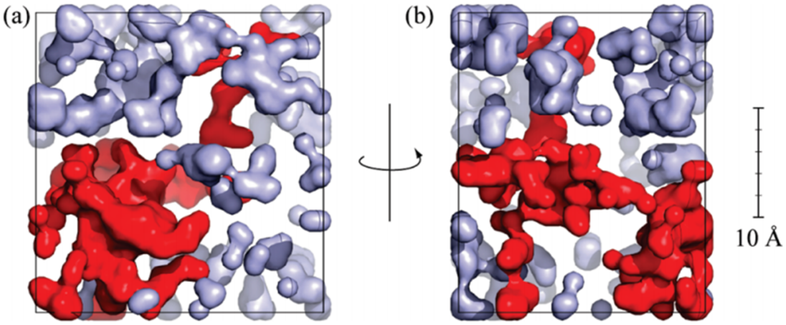


Figure 4.1: Two views of the hydrous ACC model obtained by Goodwin *et al.* using Reverse Monte-Carlo. The red surface shows the water filled percolating channel [4]. Reprinted with permission from the American Chemical Society.

using MD simulations, however, showed that the model of Goodwin *et al.* was unstable. The channels shrank during the simulation and the ions diffused towards the depleted regions to restore charge balance [175].

4.1.4 Aims and objectives

Even though a number of molecular studies have now been carried out on ACC, the impact of the protocol used to construct models of ACC on the properties and the statistical importance of the sampled configurations have never been taken into consideration. In the context of an amorphous solid, such as ACC, whether with or without water, this must always raise questions about how comprehensively the range of possible molecular configurations is being sampled. Given the experimental evidence that different forms of ACC do exist, this becomes particularly important, and may raise questions about the comparability of different modelling studies. Agreement with the limited amount of structural information is not sufficient to guarantee that the proposed structure is correct. For example, not all the proposed ACC structures that are consistent with experimental scattering data prove to be stable in subsequent molecular dynamics simulations [175].

It is therefore of considerable importance to characterise the degree of comparability with experiment (and indeed with each other) that may be expected between molecular simulation studies that use different protocols for generating the ACC structures. The purpose of this analysis is to provide a robust model of ACC based on Molecular Dynamics simulations. Such a model will be important as a benchmark for future studies involving ACC and to understand how different additives affect the structure. Additionally it will give more insights into the structure of this metastable precursor and clues on the dehydration pathway prior to the formation of a crystalline phase. To achieve this, we used four different protocols to prepare the ACC ranging from random initial structures to melted hydrated mineral structures. It is found that all protocols give good agreement with structure factors obtained from diffraction experiments. Some variations are found in the thermodynamic properties, ion coordination environments, and behaviour of water, but these variations are generally modest.

4.2 Methods

4.2.1 Simulation methods

We have used classical molecular dynamics (MD) simulations to investigate amorphous $CaCO_3.H_2O$. Force fields were selected which are known to accurately predict both the structure and energetics of crystalline $CaCO_3$ polymorphs and the thermodynamics of ion solvation, namely those of Raiteri and Gale and Demichelis *et al.* [20, 87]. Note that the use of TIP3P [180] or SPC/Fw [17] for water, respectively, is embedded within these force-fields. The two force-fields differ primarily in that the latter includes a flexible carbonate ion (and can therefore model intramolecular vibrations) in order to obtain a better comparison to solid state vibrational spectra and to capture the variation in carbonate bond lengths found in vaterite. In all other respects they predict very similar properties. Within our studies, we found no significant difference between these two force-fields, and so present the results using the Demichelis *et al.* potential [20].

All simulations were performed with the DL_POLY code (DL_POLY 4.03.4 and 2.20, and DL_POLY_Classic 1.9) [122], using the NPT ensemble at T=300 K and P=0.1 MPa implemented via a Nosé-Hoover thermostat and barostat with 0.1 ps and 1.0 ps relaxation times respectively. Orthorhombic periodic boundaries conditions were used throughout, except in the case of the monohydrocalcite structure where hexagonal boundary conditions were used [82]. The cut-off for short-range interactions was set to 9.0 Å as defined by the chosen force field [20]. Calculations were performed with electrostatic interactions treated using either the smooth particle mesh Ewald (SPME) method or the standard Ewald sum. A comparison of the calculations showed no significant difference in any calculated properties; data using SPME electrostatics are reported unless otherwise specified. The systems were usually equilibrated for 10 ns which was monitored by assessing both volume and configurational energy convergence. Then, production trajectories with a minimum time of 2 ns, though more typically ca. 5 ns, were then generated for analysis.

4.2.2 ACC preparation protocols

Four different methods of preparing the initial configuration were used. Two involved random placement of the ions, while the other two melted hydrated crystalline forms of calcium carbonate (monohydrocalcite and ikaite) followed by further steps in the case of ikaite to obtain the required $CaCO_3 : H_2O$ ratio. The structures were named ACC_{RLJ} (RLJ: Random Lennard-Jones), ACC_{RPm} (RPm: Random Packmol) for the randomly generated structure and ACC_{Ik} (Ik: Ikaite), ACC_{MHC} for the systems obtained by melting crystals. Full details are given below.

ACC_{RLJ} (Method obtained by Bushuev Y.; Warwick University): The system was initially prepared by inserting the ions and water molecules at random in a ratio of 1:3 (2880 $CaCO_3$ units and 8640 water molecules). The system was then simulated at 750 K (3 ns, NVT) with all electrostatic interactions switched off, and introducing a Ca-Ca Lennard-Jones interaction ($\sigma = 5.17$ Å, $\epsilon = 0.00674$ eV) since cation repulsion is enforced only via electrostatic interactions in our two chosen force-fields. This produces a random Lennard-Jones (RLJ) analogue of ACC. Subsequently water molecules were removed at random to obtain a composition of $CaCO_3.H_2O$, the full published force-field reinstated, and a 5 ns NPT simulation conducted at 300 K to allow the system to relax. Energy and volume were monitored over this time to confirm that relaxation was complete over the 5 ns.

ACC_{RPm}: The system was initially prepared using the Packmol package, which randomly inserts molecules or atoms within set boundaries [152]. The 486 formula units of $CaCO_3.H_2O$ were inserted with a tolerance of 2.20 Å (minimum distance between inserted particles). The system was first simulated at 3000 K (2 ns, NVT, V =40111.36 $Å^3$) then cooled down through successive steps (1 ns, ΔT = 300 K, NVT) to 300 K.

ACC_{Ik} (Method obtained by Bushuev Y.; Warwick University): A nanocrystal of ikaite (2880 formula units) was embedded in liquid water (30 000 water molecules) and simulated first at 1500 K (1 ns, NVT, V = $1.00*10^6$ $Å^3$) and then at 300 K (5 ns, NPT, P = 0.1MPa). Water molecules were then removed at random to produce a composition $CaCO_3.H_2O$, and the system again simulated at 1500 K (1 ns, NVT, volume taken from the end of the previous simulation) and at 300 K (5 ns, NPT, P = 0.1 MPa) while monitoring energy and volume to confirm relaxation was complete. The ikaite underwent a partial phase separation during the melting phase, with the consequence that the final system contained irregular pores of water, some of which spanned the simulation box; these were not found in the other three systems. This system could appropriately be described as microporous ACC.

ACC_{MHC}: A nanocrystal of monohydrocalcite (MHC) containing 576 formula units was simulated at 3000 K (2 ns, NVT, a = b = 41.90 Å, c = 30.33 Å $\alpha = \beta =$ 90.0 °, $\gamma =$ 120.0 °) and then cooled down through successive steps (1 ns, ΔT = 300 K, NVT) to 300 K while retaining the initial volume. A final equilibration simulation was conducted at 300 K (5 ns, NPT, P = 0.1MPa) while monitoring energy and volume to confirm relaxation was complete.

4.2.3 Analysis

Where needed for analysis, coordination radii were defined from approximately the first minimum in the appropriate pair distribution function: 2.8Å for Ca-O_{all}, 4.01Å for Ca-C, and 3.8 Å for $O_w - O_w$ (where O_w denotes an oxygen atom that is part of a water molecule, O_c oxygen atoms within a carbonate ion and O_{all} any oxygen atoms). To calculate free volumes and areas the Connolly surface probe method was chosen [181]. This is a method that has been commonly used in the literature to measure regions from which solvent is excluded by recording the volume accessible to a probe (see, for example its use in analysing microstructure-dependent diffusion by Zhang *et al.* (2009) and mineral-water surfaces by Zeitler *et al.* (2012) [182,183]). The Connolly free volumes and areas were calculated with the Materials Studio package using a probe with a radius of 1.575Å (corresponding to the radius of water in the potential), a Van der Waals scale factor of 1.0 and a grid interval set to 0.15 (Ultra-fine).

We have used a perfusion analysis to identify the presence of water channels. Bulk ACC simulation cells were modified by increasing the cell length in the z direction of the simulation box to convert the bulk cell into a slab separated by a vacuum gap that was subsequently filled with argon (gap width, $\Delta L_z = 50.0\text{Å}^3$, density of Ar atoms in the 'vacuum' gap, $\rho_{Argon} = 1.203\ g.cm^{-3}$, NVT, Lennard-Jones interactions for argon [184]: $\sigma = 3.405\text{Å}$, $\epsilon = 0.991\ kJ.mol^{-1}$) and all water molecules were then removed from the ACC sample. Lorenz-Berthelot mixing rules were used for all the Ar/$CaCO_3$ ϵ terms, while the σ term was chosen to be 2.3Å for all Ar-X interactions, as this corresponded approximately to the smallest distance observed in the radial distribution function between O_w and O_c.

In subsequent MD simulations (1 ns, NVT) the ingress of Ar into the ACC was used to identify potential water-filled channels accessible to the surface.

Diffusion coefficients, D, have been calculated from the long-time slope of the average mean square displacement of atoms and ions over time, following normal procedures and averaging over multiple time origins, t_0 [120]:

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{\delta < r_i^2 >_{i,t_0}}{\delta t} \tag{4.1}$$

where

$$r_i = ||r_i(t + t_0) - r_i(t_0)|| \tag{4.2}$$

and the notation $< ... >_{i,t_0}$ indicates an average over all equivalent molecules, i, and time origins, t_0. Taking an average over multiple time origins, has the advantage that more sampling is done for shorter time intervals. It was found that fitting the time interval ($0.2t_{prod}$,$0.5t_{prod}$), where t_{prod} is the total simulation time, led to a good compromise between better sampling and fitting the long time limit in 4.1. As will be shown later, and not surprisingly for an amorphous solid, the diffusion coefficients were generally very small. To obtain more information about the atom and ion mobility through ACC, we have therefore also analysed the contribution of individual molecules to the diffusion coefficient. This has been achieved by calculating the change in mean square displacement per unit time for each molecule, i, independently, thus giving a single molecule effective diffusion coefficient, D_i:

$$D_i = \frac{1}{6} \lim_{t \to \infty} \frac{\delta < r_i^2 >_{t_0}}{\delta t} \tag{4.3}$$

where the averaging is just over the time origins, t_0. The ensemble average of all these D_i values will be the normal molecular diffusion coefficient (which must be a positive value). However, individual D_i coefficients can take negative values: at any given time some molecules will be moving back towards their origin.

4.3 Results and discussion

4.3.1 Pair distribution function

Total distribution functions calculated from the simulated structures are presented in Figure 4.2, together with the experimental data of Radha *et al* [5]. It is clear from Figure 4.2 that there is good agreement between the experimental and simulated curves, and that the four different protocols lead to essentially identical total distribution functions. Such agreement is a necessary condition for a good ACC model, but is not in itself sufficient. It is only for spherical particles that the radial distribution functions give a complete specification of the thermodynamic and structural properties of the material [185]. In general, there are many configurational ensembles that map onto the same total distribution function, and so significant structural differences can still exist between systems that generate the same distribution functions. This point will be considered further in subsequent sections.

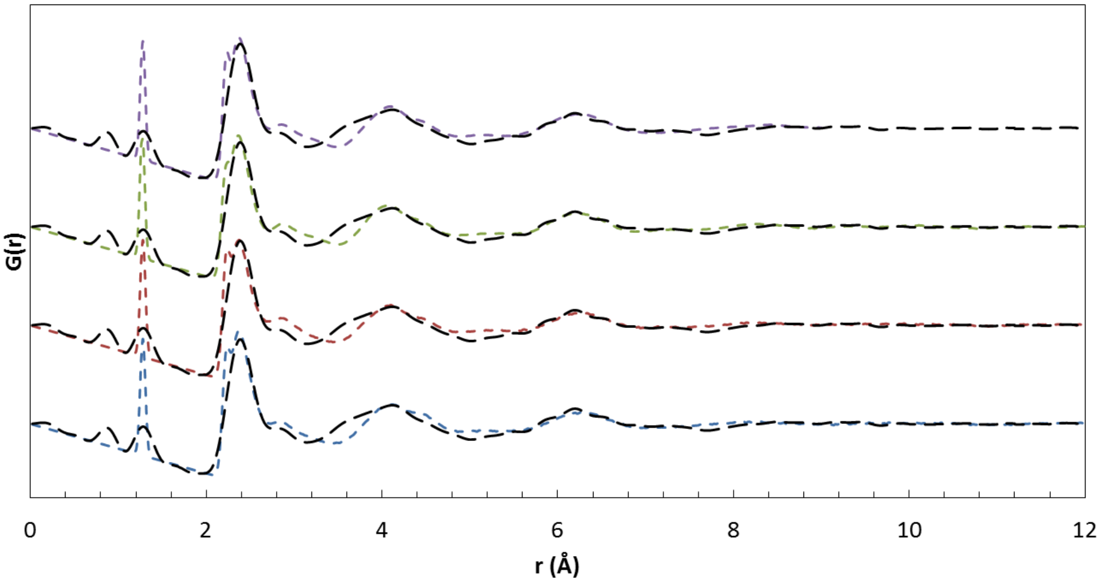


Figure 4.2: Comparison of the simulated and experimental total distribution functions; because of the similarity between the systems, four separate comparisons are presented with a vertical shift imposed to separate the curves. The curve obtained from ACC_{RPm} is blue, while red denotes ACC_{Ik}, green for ACC_{RLJ}, purple denotes ACC_{MHC} and black denotes the experimental curve obtained by Radha *et al.* [5]

The large difference between the simulations and experiment at 1.3 Å is due to a comparatively rigid $C-O_c$ bond in the force-field used for the simulations. Other small differences are apparent between the experimental and simulated curves in the region 2.5-4.0Å , and it is instructive to consider these in more detail. Because our discussion is focused on this region, we will number the peaks and troughs from the first intermolecular peak, at ca. 2.3Å. The component pair distribution functions in this region are depicted in Figure 4.3. In the total distribution function, the simulated curve shows a second peak at 2.9Å that is present more as a shoulder to the main peak in the experimental data. The experimental curve also shows a shoulder to the left of the second peak (3.5 Å) that is missing in the simulated curve, with the result that the minimum between these first two peaks shifts from 3.15 (experiment) to 3.5Å (simulation). The dominant pair contributions to this region of the total distribution function come from: O_c-Ca, which is essentially flat across this range; O_c-O_w, which shows a steady monotonic decrease in intensity; and O_c-O_c and $Ca-C$, both of which peak in the range 2.9-3.5Å. At the same time, there is very little contribution from Ca-O_w in the range 2.9-3.5Å , and only a small contribution for 3.5-4.0Å. Intriguingly, the $Ca-C$ shows two peaks, one coincident with the overstated shoulder at 2.9Å, and the other near the understated minimum at 3.15Å.

These small differences between simulation and experiment could be due to subtle deficiencies in the force-field, or to incomplete characterisation of the experimental material at the nanoscale. Recent simulation studies examining the surface organisation of

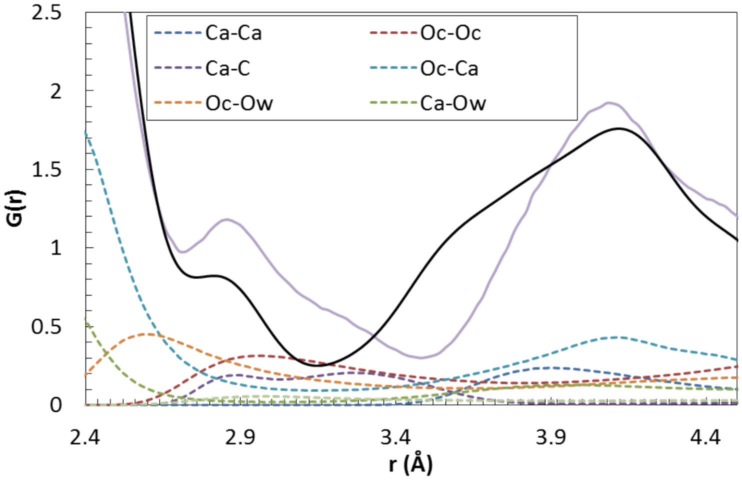


Figure 4.3: Decomposition of the calculated G(r) for ACC_{RLJ} into atom-atom pair distribution functions, g_{ij}; the smallest g_{ij} have been omitted for clarity. The black curve denotes the experimental results obtained by Radha *et al.* and the purple denotes the results obtained for ACC_{RPm}. The colour of the curve obtained for the different atom-atom pair distribution functions is shown in the legend of the graph. Data from the other ACC systems were essentially identical.

water molecules in carbonates have demonstrated that the force-fields do not perfectly replicate the X-ray reflectivity data [186], and the observations on Figure 4.3 would be consistent with a force-field that allowed $Ca - C$ contacts to become too close. On the other hand, nanoscale segregation of the water phases within ACC, impurities, or non-stoichiometry within the experimental studies cannot be ruled out. Given the potential for differences arising from the large scale structural variability of this phase, the small overall discrepancies between the simulated and experimental total distribution functions, and the coincidence of the major peak positions and heights, we conclude that the major features of ACC are well reproduced within the four series of simulations.

4.3.2 Thermodynamic properties of ACC

Potential energies and densities for the four ACC systems are shown in Table 4.1. The densities of the four simulated ACC systems fall between those of monohydrocalcite and vaterite, indicating a more densely packed structure than is found in the hydrated crystalline phase and implying that there is little space for empty pores in these models of ACC.

Comparison of the densities with experiment is difficult; there are only few results available and the composition of experimental ACCs is not often analysed. Two studies have analysed the density through SAX measurements and found values of 1.62 and 1.9 $g.cm^{-3}$, respectively [187, 188]. The main difference between the two experiments was the initial concentration of solutes, which was doubled in the second. In a further study, Faatz *et al.* also obtained a value of 1.9 $g.cm^{-3}$ using Brillouin spectroscopy for an ACC with 0.5 H_2O per formula unit [189]. The amount of water in the sample was obtained by thermo

Table 4.1: Thermodynamic properties of ACC (amorphous $CaCO_3.H_2O$) and calcium carbonate crystalline polymorphs; calculated standard deviations indicate that uncertainties are consistent with the precision quoted. The configurational energy change, ΔU_{conf}; is referred to ACC_{MHC} as the zero.

System	ρ $g.cm^{-3}$	U_{conf} $kJ.mol^{-1}$	ΔU_{conf} $kJ.mol^{-1}$
ACC_{RLJ}	2.63	-2908.62	3.38
ACC_{RPm}	2.58	-2910.45	1.55
ACC_{Ik}	2.54	-2896.44	15.56
ACC_{MHC}	2.62	-2912.00	0
Monohydrocalcite	2.37	-	-
Ikaite	1.77	-	-
Calcite	2.76	-	-
Aragonite	3.01	-	-
Vaterite	2.68	-	-

gravimetric analysis (TGA). However, Michel *et al.* used the same method as Faatz *et al.* to prepare their ACC and their TGA analysis showed a water content of 1.29 H_2O per formula unit [84]. The differences obtained within these studies, and the fact that the same experiments appear to give very different values, suggests that it is very difficult to measure the density accurately for a well-defined composition. Measurements taken at different times and in different environments give results relating to different compositions and structures. If the density measured by Faatz *et al.* is assumed to be representative, it suggests that a larger number of void spaces are present in the structure of experimental ACC than for any of the simulated structures [189]. However, such a density could also result from a more highly hydrated ACC (as for Michel *et al.* 2008) [84]. Analysis of ACC density is extremely important to better understand the subsequent crystallization process *in vivo*.

It is important to note that *in vivo* a continuous crystallization front is not always seen between ACC and calcite [89]. For instance in the sea urchin embryo's spicules, ACC is not confined to the surface as shown in Figure 4.4 *A* and *B*, but domains are observed throughout the sample [6]. It is difficult to imagine that a large change in volume could happen between the mineral phases in the bulk material. The process has been observed to occur within domains of ACC of sizes ranging from 40 to 120 nm. This suggested that the process occured throughout a particle through secondary nucleation [89]. If, as shown above, ACC with a water content of 0.5 H_2O per formula unit has a density of 1.9 $g.cm^{-3}$ and this is taken as a minimum, there could be a maximum reduction of 30% in the volume of the mineral. Such a large decrease in volume would be surprising and could lead to regions with low mineral density. These could become critical defects which would undermine the mechanical properties of the final composite. It is more likely that the vacuoles deliver an ACC particle with a density similar to the one computed in this study and to the crystalline phases. This would also have the same morphology as the final polymorph. ACC has often been observed to precipitate as nanometric spherical particles with sizes between 4 and 200 nm. In sea urchin spicules these particles, crystallize into calcite (Figure 4.4 *C*) while in Pinctada Fucata (a mollusc) they form aragonite (Figure 4.4 *D*) [7,8]. There could be slight decreases in volume, but the images obtained from electron microscopy in Figure 4.4 *C* and *D* suggest a dense structure with a small amount of pores in between the particles [8]. A similar density can be also inferred from the observation that the short range order of ACC can already show resemblence to the final polymorph and thus it is difficult to imagine large structural changes throughout the structure [10,15]. Additionally, the pair distribution function obtained here for ACC matches well to all experimental data. This suggests that the density of the experimental and *in silico* ACC should be similar.

ACC_{Ik} is the least mechanically stable configuration and shows the maximum variation in configurational energy while the other structures show configurational energies with values that vary from each other within that expected for room temperature fluctuations (RT = $2.4kJ.mol_{-1}$). For the case of ACC_{Ik} the difference is 15.56 $kJ.mol^{-1}$ which, is only 0.5% of the total energy. However it is comparable with the dissolution enthalpy of calcite (-12.5 $kJ.mol^{-1}$ experimentally [190] and -38.6 $kJ.mol^{-1}$ with the current force field [87]) and with typical hydrogen bond energies, and so cannot be considered inconsequential.

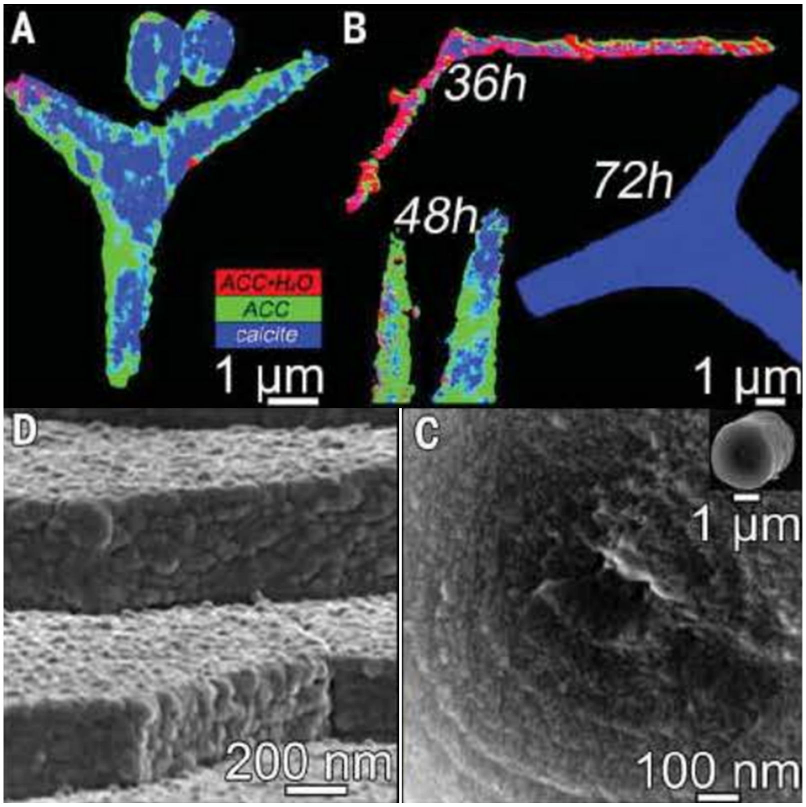


Figure 4.4: (A) Photoelectron emission microscopy component map of the sea urchin embryo spicule between 48 and 72 hours and in (B) at 36, 48 and 72 hours. In (A) and (B) the hydrated ACC is shown in red, anhydrous ACC in green and calcite in blue [6]. (C) Electron microscope micrograph of a surface of sea urchin spicules after cryo-fracture. (D) Structure of nacre tablets in Pinctada Fucata. Micrograph obtained using a field emission scanning electron microscope [7]. Permission to use the image has been kindly granted by: "The American Association for the Advancement of Science" [8].

Intriguingly, the energy differences show little correlation with densities, which indicates that the variation arise from structural differences and not just from a size-scaling of the Coulombic energy.

4.3.3 Ion coordination

The average coordination numbers for oxygen around calcium have been thoroughly analysed in the literature, and can readily be compared with our results (see Table 4.2). A more detailed analysis, in terms of the distribution of coordination numbers, is also available from simulation and is depicted in Figure 4.5. Values for the coordination numbers for our simulated ACC sits betweeen 7.43 and 7.64. This is within the range of experimental results. EXAFS of synthetic and *ex-vivo* ACC in different studies showed scattered results between 5 and 9 oxygens in the coordination sphere of calcium ions [10–16]. The distribution of $Ca\text{-}O_{all}$ in Figure 4.5 is in agreement with this spread, with Ca^{2+} ions found to have coordination numbers between 6 and 9. Our simulated data is also in good agreement with the analysis of biogenic ACC obtained from *Pyura Pachidermatina* (7.4±0.5) [14].

The simulated data (Table 4.2 and Figure 4.5) do show some statistically significant differences between the models. In particular, the model derived from ikaite, ACC_{Ik}, gives smaller $Ca\text{-}O_{all}$ and larger $Ca - C$ coordination numbers than the other three. Its distribution also shows a more even spread of $Ca\text{-}O_{all}$ coordination numbers in the range 7-8. However, as with the densities, the simulated values are actually tightly clustered when compared with the (very wide) experimental range.

The relative contributions of water and carbonate to the calcium coordination shell are

Table 4.2: Coordination number, CN(Ca.X) for element X about Ca^{2+}. The uncertainties quoted are the standard deviation of the mean. XAFS data was obtained from the following references [10–16], data from ACC from *Pyura Pachidermatina* was obtained from Levi-Kalisman *et al.* [14]

System	C_N(Ca-O_X)			C_N(Ca - C)
	O_{all}	O_W	O_C	C
ACC_{RLJ}	7.64 ± 0.01	1.71 ± 0.02	5.92 ± 0.02	5.17 ± 0.01
ACC_{RPm}	7.55 ± 0.03	1.52 ± 0.05	6.03 ± 0.05	5.20 ± 0.04
ACC_{MHC}	7.63 ± 0.02	1.58 ± 0.04	6.04 ± 0.04	5.22 ± 0.03
ACC_{Ik}	7.43 ± 0.01	1.21 ± 0.03	6.22 ± 0.02	5.54 ± 0.02
ACC (expt)	5-9, 7.4 ± 0.5	-	-	-, 4.5 ± 2
Calcite	6	0	6	6
Aragonite	9	0	9	6
Vaterite	8	0	8	6
Monohydrocalcite	8	2	6	4
Ikaite	8	6	2	1

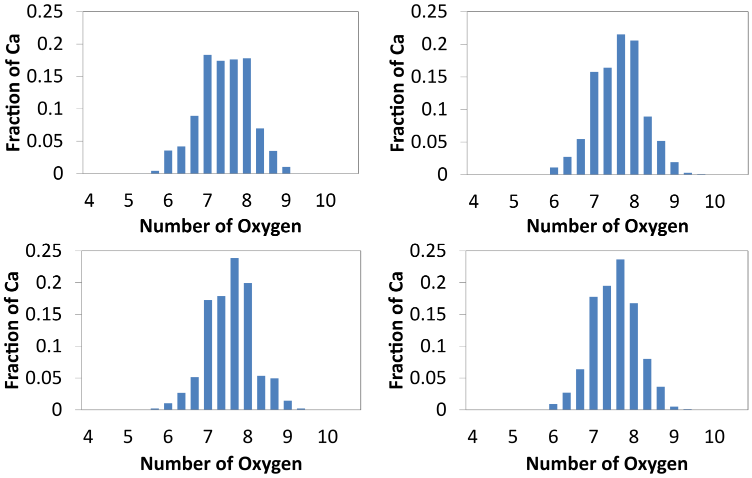


Figure 4.5: Frequency distribution of the coordination number for O_{all} atoms about Ca^{2+}, averaged across the simulation for each Ca^{2+} ion. Top row: ACC_{Ik} (left); ACC_{RLJ} (right). Bottom row: ACC_{RPm} (left); ACC_{MHC} (right).

also consistent across three of the four ACC systems, with O_w comprising 20-22% of the coordination shell. This value is actually consistent with stoichiometry: the composition of the system, $CaCO_3.H_2O$, gives a 1:3 ratio of water:carbonate oxygens; the O_c coordination is 15-16% higher than the C coordination, which indicates that 15% of the carbonate-calcium interactions are bi-dentate; this, in turn, gives a ratio for O_w:O_c of 1:(3 x 1.15), i.e. 22%, which agrees with the observed ratio of 20% to within the statistical uncertainties. The fourth model is, again, the one derived from ikaite. For this system, only 16% of the calcium coordination shell comes from water. Intriguingly, the bidentate contribution also drops to 12%. Following the same argument based on stoichiometry, this would predict 23% of the Ca^{2+} coordination shell to come from water, rather than the 16% found in reality. Thus water appears to be under-represented in the Ca^{2+} coordination environment for ACC_{Ik}. This is consistent with the formation of water micropores within ACC_{Ik} but not in the other three models.

4.3.4 Water-accessible volume and area

Connolly free volumes and areas [181], or the closely related available volumes and areas [191], are commonly-used to probe microporosity and heterogeneity within a material. The available Connolly volumes are calculated by treating atoms in the material as hard spheres and then identifying the regions where it is possible to insert a (hard sphere) probe atom without creating overlap between the material and the probe. Connolly volumes and surface areas are then calculated as the free volume (surface) accessible to the probe atom. As such, the values reported will depend on the probe radius in a way that makes quantitative comparison with experiment ambiguous. If we consider an amorphous solid, the molecules within it will not pack as perfect spheres, but will form interstices of irregular shape. As a result, if we try to obtain a value of the available volume occupied by water molecules using a spherical probe with the Van der Waals radius of water, 1.575Å (shown in Table 4.3) we obtain values that are considerably lower than the average volume of a molecule in liquid water (30.0Å^3), or indeed of that reported for a water molecule within monohydrocalcite [87]: 14.86 Å^3. On the other hand, using a probe radius of zero will overestimate the volume (area), as it overlooks the inefficient packing of spheres in 3D space that happens with real atoms and molecules. In the case of ACC, this uncertainty will be accentuated by strong electrostatic repulsions (O_c-O_w) and attractions (Ca-O_w), which undermine the concept of using a unique probe radius. As a result there is no unique and general way of mapping available volume onto experimental measures. Nevertheless, these measures of volumes and surface areas are extremely useful for comparing theoretical models (where probe definitions can be used consistently) and so have been used here to assess the effects of the four ACC preparation protocols. In particular, it is possible to calculate the Connolly volumes accessible to the probe for a series of probe radii to gain much greater information about the size-dependence of pore space within the material.

The Connolly volumes accessible to the probe have been calculated for all four ACC structures. The calculation has been done for the ACC as formed (to give a measure of the native porosity of the material) and with the water molecules removed (to give a measure of the volume occupied by water). We shall refer to this as 'dehydrated ACC' as no relaxation to a stable anhydrous ACC has been allowed. The results are given in Table 4.3. There is a large variation in the percentage volume occupied by the water molecules between the ACC samples when using a probe radius of 1.575Å. ACC_{Ik} clearly shows the highest water-occupied volume with 13.62%, while the lowest value is observed for ACC_{MHC} with 2.85%.

For the 'dehydrated' ACC, we have also calculated the Connolly volume accessible to the probe across a wide range of probe radii (Figure 4.6). The probe radius dependence follows the trends in the density identified above. Three of the systems (ACC_{RLJ}, ACC_{RPm} and ACC_{MHC}) follow a similar curve and show very similar magnitudes. The ACC_{RPm} curve is shifted up by about 1.2%, which is consistent with its density being about 1% lower than the other two (Table 4.1). By contrast, the ACC_{Ik} system shows consistently higher probe-accessible volumes (about 6% at smaller probe radii, again consistent with its lower density). Much more significantly, however, it shows a much slower decay as the probe radius increases. This is indicative of larger pore sizes (the probe excluded volume effects are confined to the pore surface) and is entirely consistent with the formation of nanoscale channels in the ikaite-derived model.

The maximum Connolly accessible volumes recorded in Figure 4.6 are about 30%, and correspond to a volume of 20.23Å^3 per water molecule. This is significantly larger than the value of 14.86Å^3 calculated for monohydrocalcite by Raiteri and Gale (2010) [87]. However this should be considered as an upper bound. As mentioned before, this value will include volumes that are not solely occupied by water molecules, such as small cavities present within the structure due to the inefficient packing of the ions. When the probe radius is increased to 1.575Å (the approximate van der Waals radius of a water molecule in our model) the accessible volume drops to 2.4Å^3 per water molecule for three of the systems, and 10Å^3 for ACC_{Ik}, as shown in Figure 4.6. For comparison, the corresponding calculation for ACC_{MHC} with the water removed yields an available volume of just 0.055Å^3. Given the discussion of liquid water free volume above, this indicates that all four ACC systems exhibit a degree of nanoporosity but that this level is again higher for ACC_{Ik}. As noted in the methods section, ACC_{Ik} could reasonably be considered microporous.

As all the structures were prepared with the same stoichiometry, the difference must arise from the configurations that the water molecules adopt within the different structures. The accessible surface area can be used to investigate some of these differences. Table 4.3 shows that the surface area available to the probe is approximately equal for all the models for a probe size equivalent to a water molecule, although the corresponding accessible volume is much larger for ACC_{Ik} than for the other cases. The smaller surface to volume

Table 4.3: Free Connolly volume and surface area within the system for the four ACC models using a probe radius of 1.575 Å

	ACC_{RLJ}	ACC_{RPm}	ACC_{MHC}	ACC_{Ik}
Percentage free volume accessible to the probe after water is removed from structure	2.69	5.36	3.57	13.621
Free volume (Å^3) previously occupied by water accessible to the probe per formula unit	1.99 ± 0.02	4.04 ± 0.11	2.66 ± 0.01	10.5 ± 0.05
True free volume per formula unit (Å^3) for ACC ($CaCO_3.H_2O$)	0.01 ± 0.00	0.00 ± 0.00	0.05 ± 0.02	0.01 ± 0.01
Area (Å^2) of the free surface accessible to the probe when water is removed from the structure	4.06 ± 0.06	7.81 ± 0.11	5.88 ± 0.03	7.45 ± 0.03

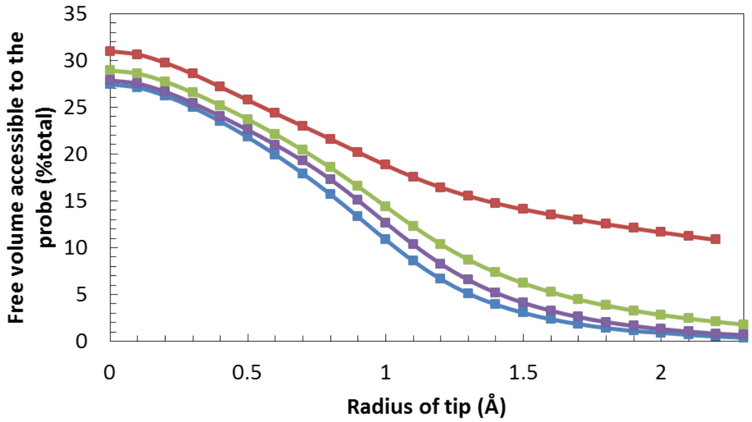


Figure 4.6: Change in free Connolly volume accessible to the probe (as a percentage of the total volume) after water is removed from the structure as a function of the probe radius. The blue curve refers to ACC_{RLJ}, red to ACC_{Ik}, green to ACC_{RPm} and purple to ACC_{MHC}

ratio for ACC_{Ik} suggests that the water molecules in that system are more segregated from the ionic framework, forming water channels, (Here 'channel' means a pore large enough to allow the translational movement of a water molecule.). This gives some of the water in ACC_{Ik} properties closer to that of a liquid. On the other hand, ACC_{MHC}, ACC_{RLJ} and ACC_{RPm} models probably form small, individual pores within the structure. This can be seen in Figure 4.7 where, in the ACC_{Ik} model, the formation of a water layer (in blue) at the centre of the periodic box is clearly visible and a percolating channel appears present whereas the other three structures show a much more homogeneous water distribution throughout the structure.

The change in the total accessible volume occupied by water along the trajectory of the simulation has also been analysed. The total accessible volume changes by approximately 5% over a relatively small amount of time (approximately 0.1 ns) indicating that small

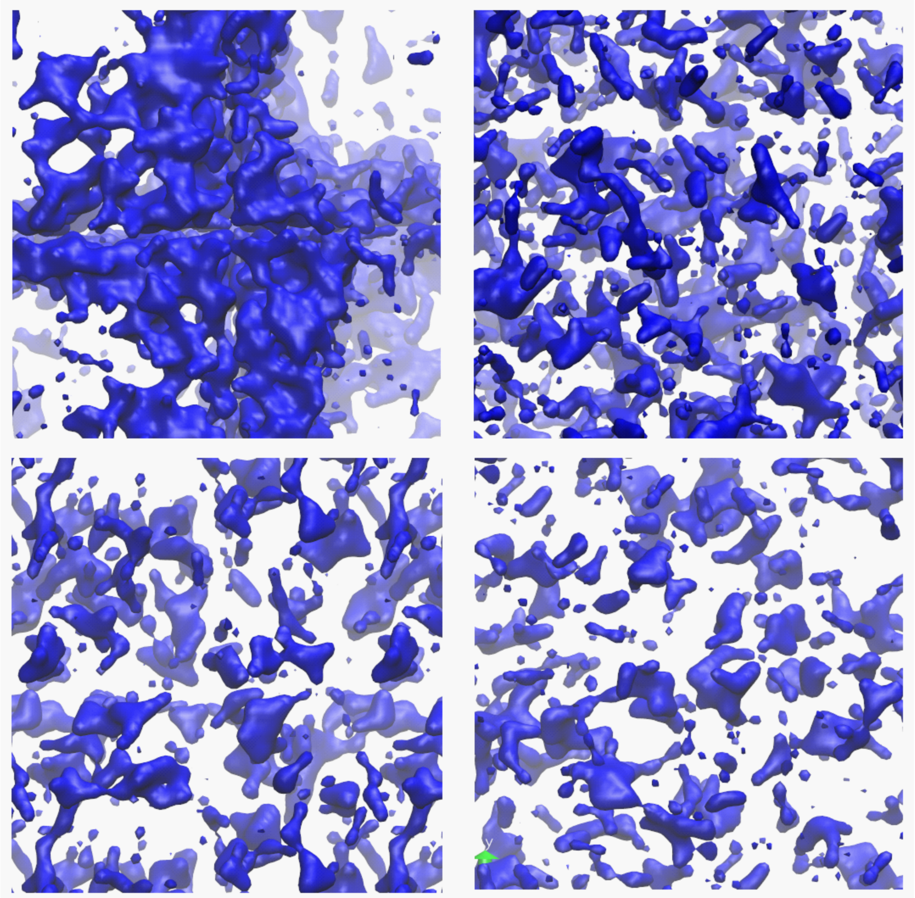


Figure 4.7: Representation of the ACC structure obtained using a surface representation of the O_w from the VMD visualiser using the QuickSurf drawing method (radius scale = 1.1, density isovalue = 1.0, grid spacing = 1.0) to show only the water molecules (in blue in the image). Top row: ACC_{Ik} (left); ACC_{RLJ} (right). Bottom row: ACC_{RPm} (left); ACC_{MHC} (right)

scale rearrangements within the system are continually occurring. This mechanism could lead to possible dehydration pathways such as the transient formation of channels.

4.3.5 Water channels and perfusion of argon

As will be shown below, the $CaCO_3$ framework shows little mobility during any of our simulations. As a result, it is possible to estimate the volume accessible to water by calculating the average 3-dimensional density distribution of water over the simulation; this is akin to visualizing the trajectory with time-lapse photography. The results are depicted in Figure 4.8 (*a*). The average volume occupied by water molecules during a simulation shows chains of water molecules that pervade the whole structure.

To further explore the possibility that these chains could potentially be percolating channels, and provide a transport network for water within the ACC, anhydrous ACC films were created by removing water from the bulk ACC (see methods section), and simulated under an Ar atmosphere. Ar-ACC parameters were modified to give Ar an apparent size within the ACC that mimicked water. Perfusion of Ar atoms into and potentially through the dehydrated ACC film would then identify channels through which water transport was sterically feasible. A snapshot of the distribution of Ar atoms during these perfusion simulations is given in Figure 4.8 (*b*). The microporous system, ACC_{Ik}, showed the presence of pores that percolated through the system. Pores were also observed for all the other structures. However these were all relatively short, with the longest pore penetrating only 13-15Å into the slab. This is far less than the width of the slab of 33Å , and demonstrates the absence of stable percolating channels within these three ACC structures. Coupled with the results detailed for Figure 4.8 (a), this suggests that the water movement is a dynamic process where the water is not segregated from the ions but can jump from one site to the next. The ACC_{Ik} structure was the only one to show fully connected channels of water. As shown in the section discussing the thermodynamic properties of ACC, this structure had a higher configurational energy compared to the other three ACC structures and is therefore likely to exist (if at all) as only a small fraction of the population of the ensemble of structures present in ACC and we would not expect it to contribute significantly to the average structural data. This is in contrast to the results of Goodwin *et al* [4]. Thus it would be difficult for channels to form, but their presence is potentially possible as shown in this work and previous analysis [179]. A recent study by Bushuev showed that percolating chains of water molecules would be present for composition of ACC containing more than 0.8 H_2O per formula units [179]. However it is difficult for these to be the main route for water expulsion during formation of transient anhydrous ACC. The percolating chains described by Bushuev *et al.* will probably have approximately the same diffusivity as the other water molecules in the structure as will be shown in the next section.

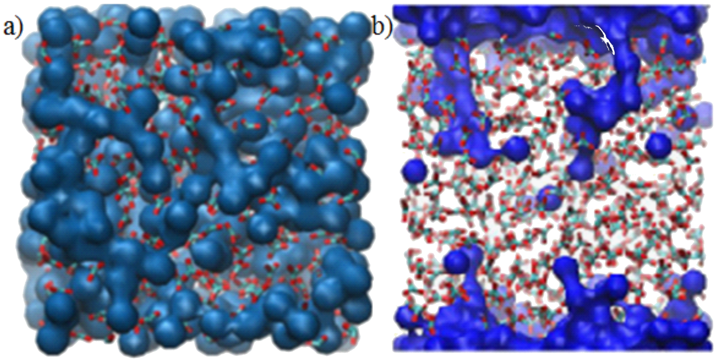


Figure 4.8: (a) Average density of water molecules during a simulation of ACC_{RPm}. Key: Ar is represented by blue spheres with radius 2.3Å (the Lennard-Jones radius); (b) Distribution of Ar atoms within the ACC_{RPm} model during MD perfusion simulations. Single snapshot showing channels accessible by Ar from the dehydrated ACC surface; red (O) and cyan (C) lines depict the bonds within carbonate ions; the water molecules are represented by cyan spheres.

4.3.6 Atom and ion mobility

The dynamics of a system are likely to depend strongly on the material structure. Analysing the mobility of the atoms and ions through ACC should therefore be a sensitive test of comparability between the different ACC models. Diffusion coefficients have been calculated from the slope of the ionic/atomic mean square displacements with respect to time (the D defined in the methods section) and results are presented in Table 4.4. From these values we can see that the only water diffusion coefficients large enough to suggest significant water mobility are those for ACC_{Ik}. Moreover, the distribution (see also the insert in Figure 4.9) shows that this value of the molecular diffusion coefficient is dominated by a comparatively small population of mobile water molecules with diffusion coefficients in the range 10^{-11} - 10^{-10} m^2s^{-1}. This suggests that these water molecules inhabit a different

Table 4.4: Diffusion coefficients for the components of ACC

System	10^{-15}D / m^2s^{-1}		
	Ca^{2+}	CO_3^{2-}	H_2O
ACC_{RLJ}	5.1	5.5	42
ACC_{RPm}	13	14	110
ACC_{MHC}	6.4	5.8	48
ACC_{Ik}	32	33	1000

part of the amorphous structure to most of the molecules – the micropores. Water within these micropores will be significantly more mobile than water within an ionic glass, and so gives an average diffusion coefficient that is much larger than for water in the non-microporous ACC systems.

This is reminiscent of the experimental results obtained by Michel *et al.* and Reeder *et al.* using proton NMR to analyse synthetic ACC [84, 85]. In that work half of the water molecules were largely immobile and were labelled as structural while the rest were capable of restricted motion. If we look at the single molecule diffusion coefficients for the water molecules defined using H_w (Figure 4.10) a much wider spread in the data is found, more in accordance with the proton NMR. This spread of values can be ascribed to rotational motion of the water molecule, giving comparatively large motions of the hydrogen atoms. Thus some of the water in ACC is translationally free, moving down channels, whereas some is rotationally free, as suggested by the NMR, and some is fully incorporated into

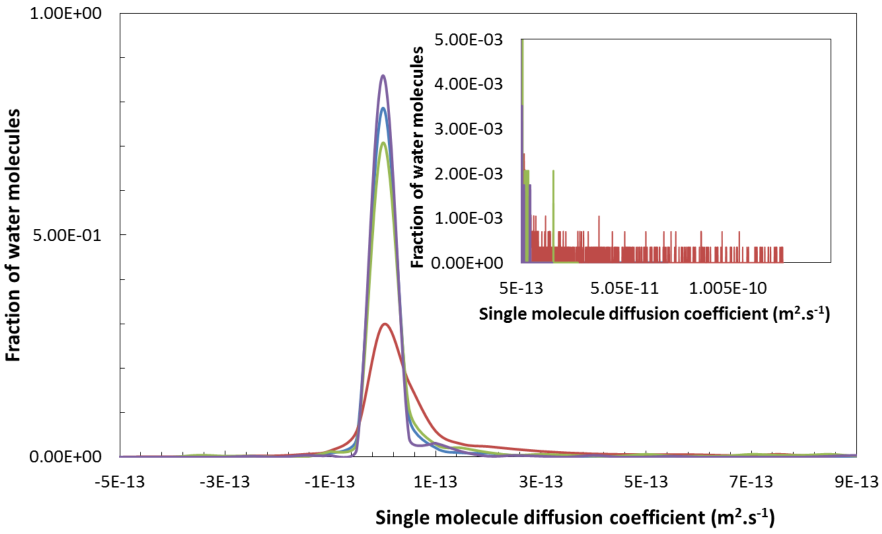


Figure 4.9: Distribution of single molecule diffusion coefficients D_i (see the method Section 4.2.3, and Equation 4.3) for the oxygen of the water molecules for the different models. The blue, green, red and purple curves represent ACC_{RLJ}, ACC_{RPm}, ACC_{Ik} and ACC_{MHC} respectively.

the structure of the amorphous solid.

Michel *et al.* also found that the mobile water was associated with the carbonate ion and the structural water was inferred to be closer to the Ca^{2+} ions. This contrasts with another study looking at biogenic ACC [85]. Here the authors did not find mobile water close to the carbonate ions, suggesting segregation of the water molecules within the structure forming channels and pores. Again, when the pair distributions functions were compared, no clear differences were observed between the synthetic and biogenic samples. They hypothesized that in synthetic ACC the water mobility permeated the structure, which could aid ionic transport and ACC restructuring, while in biogenic ACC the water

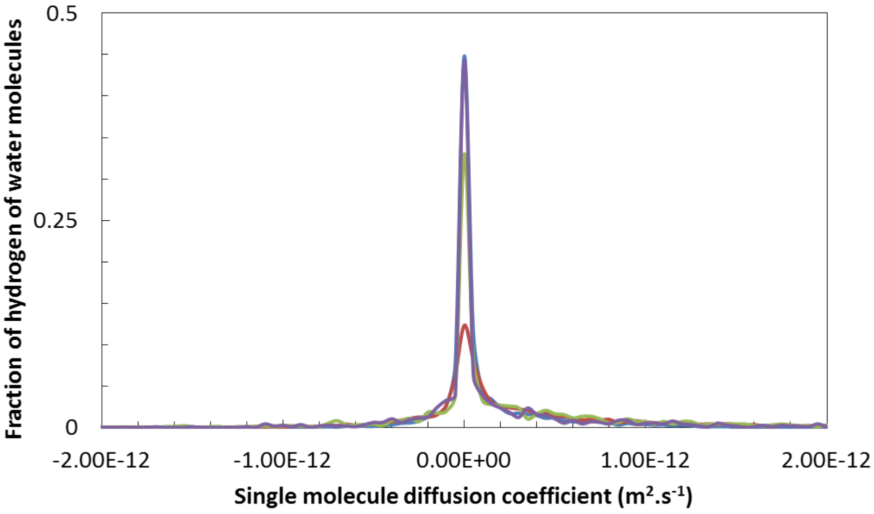


Figure 4.10: Distribution of single molecule diffusion coefficient, D_i (4.2.3, and Equation 4.3) for the hydrogens of the water molecules for the different models. The blue, green, red and purple curves represent ACC_{RLJ}, ACC_{RPm}, ACC_{Ik} and ACC_{MHC} respectively.

and carbonate are decoupled leading to a more rigid structure which could explain the increased stability of the sample. However, although our model ACC_{Ik} shows the most segregation of water within the structure, it does not give the slowest diffusivity for the ions. While the more mobile water will be more likely to leave the system during dehydration, full dehydration is likely to involve further structural changes.

4.4 Conclusion

In this Chapter we have examined the extent to which different protocols for creating hydrated amorphous calcium carbonate (ACC) for simulations can give rise to observable differences in the calculated physical properties. Experimental studies do suggest that different forms of ACC exist. If this is mimicked in simulations, where a much greater degree of control and characterisation of the nanostructure can be exerted, it would have important ramifications for our ability to compare and interpret simulations of amorphous calcium carbonate.

Four different protocols have been used in this work, starting from both hydrated crystalline phases and random structures and employing a wide range of temperatures during the initialisation stages. All four gave very similar pair distribution functions for the ACC phase. Associated total distribution functions and ion coordination numbers were also in good agreement with available experimental data for all four simulated ACC systems. Indeed, none of the simulated ACC structures were inconsistent with the available experimental data for real ACC materials.

Within this envelope of agreement with experiment, however, some significant differences in the physical properties of the ACC models were found to exist. While three of the models gave similar behaviour for all properties examined, the protocol based on melting and then dehydrating ikaite crystals did introduce important differences in the ACC structure. In particular, it generated a partial phase separation that gave rise to water-filled micropores within the ACC. This led to a lower overall density and a faster diffusion the water molecules. It did not, however, increase the mobility of Ca^{2+} and CO_3^{2-} ions; nor did it substantially change the coordination environment of the Ca^{2+}. Interestingly, the phase separation of the water molecule did lead to a structure which was mechanically less stable when compared to the other more homogeneous models. The formation of channels is potentially possible, as also observed in our models. However, as the configurational energy is higher (i.e. less stable) this structure would have a low weight and in the absence of external factors ACC will probably express a more homogeneous configuration. Thus the process of water expulsion must happen through a different mechanism. If we consider the results of Raiteri and Gale, the surface of ACC would be more hydrated and could dissolve until reaching the 'anhydrous' core [87]. This would then be able to crystallize via either a dissolution-precipition mechanism or solid-state transformation. However to verify this, more analysis of the ACC surface in contact with water must be performed.

We conclude that simulated ACC mimics the behaviour of real ACC in giving rise to several slightly different amorphous forms. In general, the variation between the simulated systems is less than that observed between different experimental studies at the same composition of $CaCO_3.H_2O$, and ensures that it is reasonable to expect overall comparability between different ACC simulations. Greater experimental resolution of the nanoscale structure of real ACC materials would, however, be helpful in developing more definitive models of ACC.

4.5 Future work

Additional work on ACC have been carried has this work was getting published. As mentioned, Bushuev *et al.* analysed the thermodynamics of ACC showing a minimum in the free energy for 1-6 water molecules per formula units [179]. The dehydration process was also examined remarking that the mobility of water molecules within ACC decreased during dehydration thus hindering the process [179].

This ACC protocols can now be used to prepare reliable surfaces of ACC. They could then be used to analyse water structure and compare it to different models, and to study the binding of organic molecules and inorganic ions. The results could help shed more light on the specific interactions responsible for the formation of different morphologies and polymorph selection. The next chapter will analyse some of these interactions by investigating the binding of four different amino acids on both anhydrous and hydrous ACC surfaces.

Rare events methods such as Metadynamics could be used to analyse the crystallization mechanism(s). Simulations of ACC surfaces in contact with water and selections of solutes could give additional details on how these affect the crystallization process.

Chapter 5

Interaction of ACC with amino acids - Can everything stick to its surface?

5.1 Background

Living organisms have developed the ability to precipitate minerals with morphologies controlled at the molecular level and properties finely tuned for their task. While their capability is controlled by a number of parameters, two factors seem to be highly important to the process; the initial precipitation of an amorphous precursor and the interactions between the mineral phase and organic molecules. Nucleation of precursor amorphous phases occurs throughout the animal kingdom [12, 58]. Thus, understanding the effect of organic molecules on the structure of this intermediate phase and on its potential energy landscape could have an important impact on our grasp of biomineralization. The results could then be used to design new manufacturing processes for advanced materials.

5.1.1 Experimental observations of the effect of organic molecules on nucleation and growth

Interactions between ACC and organic molecules are essential in the formation of the required crystalline polymorph [12, 192, 193]. While studying molluscs, both Falini *et al.* [66] and Belcher *et al.* [194] showed that polymorph selection can be controlled by macromolecules associated with the prismatic and nacre layer of molluscs. Later, when an amorphous precipitate started to be observed *in vivo*, these results suggested that macromolecules influenced polymorph selection by modifying the structure of the amorphous precursor [63, 88]. Confirmation was obtained when the amorphous precipitate in both aragonitic-forming molluscs [10] and calcitic-forming sea urchins [15] showed structural patterns that matched the final polymorph.

Stabilisation of this amorphous phase is extremely important in living organisms. First of all it is an easily accessible source of ions and secondly it is essential in controlling both the location and the timeframe of crystallization [12]. Interaction of the mineral phase with organic molecules, such as proteins and saccharides, is believed to be essential in the process, even though physical constrains have also been observed to be important [167, 192, 195]. Following the observation of a large number of acidic domains in organic molecules involved in biomineralization, most of the research focussed on these [63]. Early investigations showed that in molluscs, amorphous calcium carbonate is found in the presence of proteins containing substantial amounts of glutamic acid and glycine [196, 197]. On the other hand, calcite crystals seemed to be associated with proteins having long sequences of aspartic acid residues [196]. Interestingly, however, *in vitro* precipitation of calcium carbonate solutions, using poly-aspartic acid as a proxy of organic molecules, stabilized a liquid amorphous calcium carbonate phase in solution [91]. The difference between the two observations resides in the 3D extracellular matrix present *in vivo*. For instance, the presence of chitin and silk as a scaffold, in addition to the correct protein, is essential in the formation of aragonite [68]. On this scaffold, poly-aspartic acid led 60 % of the time to higher aragonite yields compared to calcite. This result show that acidic proteins or polymer can affect both nucleation and growth, and that living organisms can modify the function of the organic molecule depending on its location.

Contrary to previous belief, not only acidic domains are able to interact with the precipitating mineral phase. In a recent study Cantaert *et al.* showed that the same results obtained with poly-aspartic acid in solution could be replicated using a basic organic molecule: poly(allylamine hydrochloride) [113, 115]. In another study, Tavafoghi and Cerruti analysed the effect of glutamic acid and arginine on the precipitation of hydroxyapatite from solution [161]. Arginine was better in delaying the precipitation of HAP [161]. Additionally its incorporation within the structure was much more pronounced [161]. Thus, it appears that both basic and acidic residues can have an impact on the precipitation of the mineral phase. The latter study also shows that, in order to modify the potential energy landscape of nucleation and growth, proteins or large molecule assemblies are not essential and similar effects can be obtained using more fundamental molecules. This is extremely important as amino acids are much easier to understand (lack of the complex tertiary sequence) and cheaper to manufacture and thus, can have a much larger impact on the respective industry.

5.1.2 Computational studies of organic molecules binding to ions and mineral surfaces

Important fundamental mechanisms can be extracted from molecular simulations, which could lead to an increase level of control over the system via a better understanding of its principles. For instance, the chicken protein ovocleidin-17 has been extensively studied by Freeman *et al* [9, 163, 198, 199]. In their analysis, it was found that ovocleidin-17 can act as a catalyzer to the formation of calcite from ACC [198]. Using metadynamics it was possible to recover the free energy landscape as a function of the Q_4 Steinhardt order parameter [198]. Ovocleidin-17 lowered the barrier to the crystallization of calcite and, as it accomplished its purpose, detached from the crystalline nanoparticle ready for another cycle [198]. The interaction of this molecule with anhydrous ACC and calcite were then examined in more detail. The binding energy of the protein was found to be stronger on the amorphous phase [9, 163, 199]. Interactions on the amorphous surface occur mainly through basic groups in the random coil section of the protein, as shown in Figure 5.1. Arginines are highly involved in the interaction, through the guanidine groups (Figure 5.1). Serine and glycine as well were observed to be important. The latter were even observed to partly incorporate within the mineral surface [9]. All of these studies used the AMBER standard force field to model the intermolecular interactions of the organic molecule, but other force fields do exist.

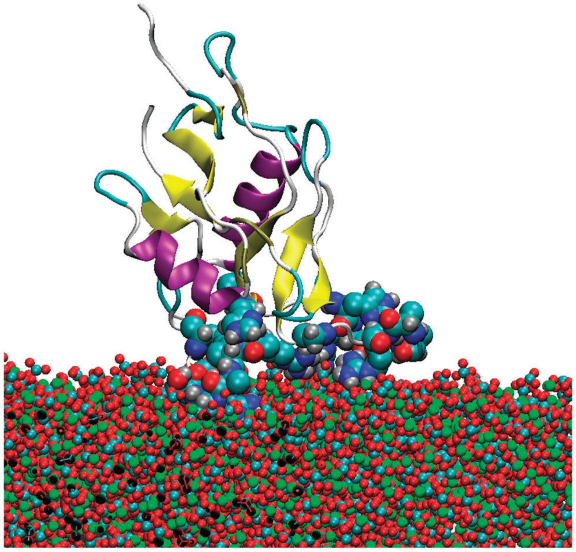


Figure 5.1: Image of ovocleidin-17 on the surface of amorphous calcium carbonate. The nitrogen atoms are shown in dark blue, the oxygens in red, the carbons in light blue and the calciums in green. It is clear from the image that arginine is the main amino acid involved in the binding of the protein [9]. The image was reprinted from Freeman *et al.* (2015) with permission from the Royal Society of Chemistry.

Raiteri *et al.* used the CHARMM force field to analyse the interaction of calcium and carbonate ions, their ion pair, amorphous nanoparticles and the {104} calcite surface with citrate, aspartate and acetate. [200]. The simulations showed that the free energy of binding of acetate and aspartate to the amorphous surface was very similar. Aspartate, however, showed a more favourable pairing to Ca^{2+} and it was suggested that the formation of this complex allowed aspartate to bind to intermediate phases and affect their stability [200]. In another study, Saharay and Kirkpatrick used metadynamics to obtain the free energy curve of the interaction of functional groups representing different amino acids [201]. The functional groups were modelled using guanidium, acetate, acetic acid and ethanol. As the strongest binding energies were found for the guanidium and acetate groups, it was suggested that these acted as nucleation sites for the crystallization of calcium carbonate [201]. Taking into consideration the previous results obtained by Freeman *et al.* studying ovocleidin-17, this appears highly probable from an *in silico* perspective. However, the binding of organic molecules to ion pairs must be considerably different from a surface, as shown by Raiteri *et al* [200]. The surface could form water layers, which could impede the binding, or modify it. Raiteri *et al* [200] did not look at positively charged functional groups in their study, thus it is important to analyse the binding of these and compare them to negatively charged ones in a self-consistent study.

5.1.3 Aims and objectives

The aim of this study is to understand how different amino acids bind to an ACC surface and the effect of the binding on the interface. As organic molecules are known to affect the structure of ACC, MD simulations can extract the atomistic details of the process.

Here we analyse the binding of glutamic acid, aspartic acid, glycine and arginine to both hydrous and anhydrous ACC surfaces using classical MD simulations. These amino acids were observed in a large number of proteins involved in biomineralization and more recently in contact with the highly stable ACC produced by earthworms [202]. Additionally, they have different properties in solution. Two of the amino acids are acidic (glutamic acid and aspartic acid), one is basic (arginine) and the fourth is a neutral amino acid (glycine). MD was chosen to obtain an atomistic view of the process and to investigate the dynamics of the system. The results show that all amino acids could potentially bind to the surface. The favoured mode of interaction observed was hydrogen bonding between either the guanidine or amine groups. Furthermore, the amino acids modified the local dynamics of the water molecules in contact with the amorphous surface. The implications of the results with regards to biomineralization will be discussed.

5.2 Methods

5.2.1 Preparation of system

The ACC slabs were prepared using the method described in Section 4.2.2, which leads to a structure fitting well to the experimental pair distribution function and with a comparable coordination number for the different ions in the system [203]. Briefly, the ions were randomly packed in a cube by using the program PACKMOL (the minimum distance between elements was set to 2.2 Å) [152]. Two different configurations were prepared. One contained 294 formula units of $CaCO_3(ACC_1)$ and the second 294 formula units of

$CaCO_3.H_2O$ (ACC_2). The configurations were then melted for 1 ns at 3000 K using a timestep of 1fs in an NVT ensemble using DL_POLY_classic version 1.9 [204]. A Nosé-Hoover thermostat with a 0.1 ps relaxation time was used to control the temperature of the simulation. Then, using the same simulation parameters, the configuration was cooled to room temperature (300 K) in steps of 300 K. The ensemble was then changed to NPT using a Nosé-Hoover barostat (to control the pressure of the simulation, set at 0.01 atm) and thermostat. The relaxation times were set to 1.0 and 0.1 ps, respectively. The simulation was continued until the configurational energy, volume and temperature reached a stable state. The statistical data presented are extracted from 5 ns of additional simulation time.

To prepare the complete systems, including the ACC slabs, the water and amino acids, PACKMOL was also used [152]. The minimum distance between the different elements of the system was set to 2.2 Å. The distance between the amino acids and the ACC surface was set to a maximum of 5.0 Å. When the amino acid sampled was charged, a counter ion was inserted in the simulation box; either Cl^- (for arginine) or Na^+ (for glutamic acid and aspartic acid).

5.2.2 System simulation

For all the following work we used the force field presented by Raiteri and Gale (2010) [87] to model the ACC. As suggested in the specification of the carbonate force field, water molecules were treated using the SPC/F model [17].

The amino acids intramolecular interactions were modelled using the standard AMBER force field set, more specifically the ff12SB form [154]. As discussed in the introduction, other force fields are available in the literature, however none of them have been demonstrated to be more accurate than the one chosen here [205,206]. Additionally, the AMBER force field has been used in a large number of previous analyses of the calcium carbonate system, allowing for a consistent comparison of the results [9,207,208].

Cross terms between the organic and mineral phase were fitted using the Schroeder method described by Freeman *et al* [21]. The intermolecular interactions of the organic molecules with water were obtained by using the standard Lorentz-Berthelot rules.

Four different configurations of each amino acid, chosen randomly, were analysed for both the hydrous and anhydrous ACC slab. In all the different configurations the amino acid was tethered using a harmonic potential (k=2.5 $kJ.mol^{-1}$). All the simulations were equilibrated using a NVT ensemble implemented via a Nosé-Hoover thermostat (relaxation time of 0.1 ps). The timestep was set at 1fs and the simulation was continued for 10 ns. The system never fully equilibrated, even though it was slowly converging. The ensemble was therefore set to NPT using a Nosé-Hoover thermostat and barostat (relaxation time of 0.1 ps and 1 ps, respectively) to improve convergence. Most simulations were continued for approximately 30 ns. A number of the simulations including the amino acids, failed to reach a satisfactory convergence after 30 ns. The most stable configurations were then simulated without tether for an additional 15 ns, where the systems finally reached convergence. The results presented are from the last 5 ns of these simulations. For all the above simulations the electrostatic interactions were treated using the Ewald summation method and the cut off for the intramolecular interactions was set to 9 Å.

5.2.3 Calculation of amino acid binding energies to ACC

The binding energies of the amino acids were calculated using the following equation:

$$E_{binding} = E_{allsystem} - (E_{water-carbonate} + E_{water-molecule-ion}) + E_{water} \tag{5.1}$$

where $E_{allsystem}$ is the configurational energy of the system including the amino acid, the water, the ACC and the counter ion. $E_{water-carbonate}$ contains only the water and the calcium carbonate ions, $E_{water-molecule-ion}$ is a simulation of the water molecules, the amino acid and the counter ion, and E_{water} includes only the water molecules.

5.2.4 Analysis of residence times

The residence times of the water molecules were calculated by counting the number of consecutive time steps that a water molecule would stay within the first coordination shell of the element of interest. If a water molecule left this coordination shell for more than 2 ps then the count would be reset. Thus, the residence time does not give a measure of the time it takes for a water molecule to transition from a "bound" state to free water. Instead, it measures the time that a water molecule exists at a given place, which is here defined as the first coordination shell of the carbonate's oxygens.

Compared to the Pavese *et al.* [178] model, previously used extensively for the simulation of calcium carbonate surfaces, the force field model of Raiteri and Gale [87] shows some ion dissolution [2]. Thus, the ACC surface is rough, and as shown above, water molecules extensively penetrate inside it during equilibration. The residence time of the water molecules will be highly dependent upon their depth within the structure and their interactions with the ionic framework. Some of the water molecules will have become embedded within the ionic structure and will have infinite residence times. The decision on which water molecules to include in the calculations of the residence times will have an impact on the results and moreover is somewhat arbitrary. Here, in order to discriminate between the 'structural' and 'liquid' water molecules, the number of O_c (carbonate oxygens) present in the coordination shell of O_w (water oxygens) was chosen. Since in bulk hydrous ACC the average coordination number is five, it was decided to use four as the limit. If a water molecule had more than four O_c in its coordination shell it would not be used in the calculations. The analysis was also conducted using as a limit coordination numbers of three and five O_c without affecting the results. The residence time of water molecules in the coordination shell of the carbonate oxygens was calculated on both the surface of ACC_1 and ACC_2 with and without amino acids.

5.3 Results

5.3.1 Structure of bulk ACC and surface

The bulk ACCs were in good agreement with the experimentally obtained pair distribution function (PDF)(Figure 5.2). The PDF obtained for both ACC_1 and ACC_2 was approximately the same, showing that water has only a small effect on the overall structure of ACC, as recently shown by Bushuev *et al* [179].

When water molecules were added to the ACC slab, a diffuse interface formed. In the case of ACC_1, which was anhydrous, Figure 5.3 shows that water penetrated approximately

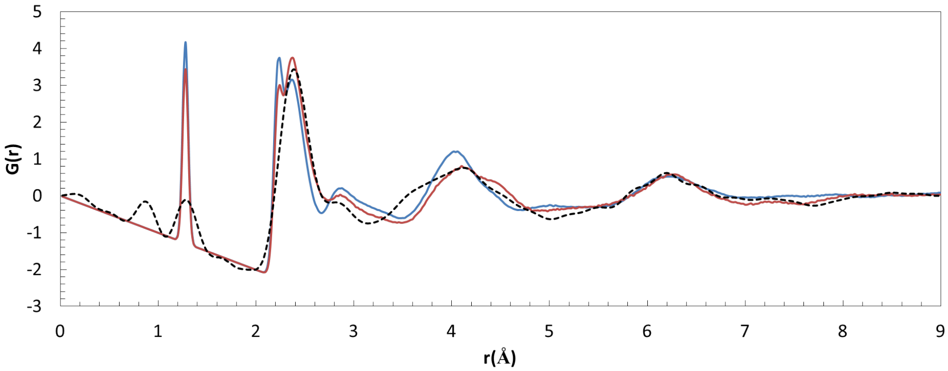


Figure 5.2: Pair distribution function of the simulated ACC discussed here. The blue line represents ACC_1, the red ACC_2 while the black line is an experimental result obtained by Radha *et al* [5].

10 Å within the structure. Both ACC_1 and ACC_2 showed dissolution of the mineral phase. Compared to the crystalline surfaces [9], the ACC surfaces do not structure water molecules as strongly. The water layers are diffuse and do not differ much from the bulk structure. In the case of calcite, the first water layer can be four times as dense as the bulk [9]. By contrast, here the first peak is 1.2 times the bulk value. This is important as incoming molecules will not have to compete with water molecules as strongly bound to the surface, which means that the free energy barrier associated with the binding process will be smaller compared to a crystalline surface.

The diffusion coefficients of the water molecules and the ions along the z direction of the system have also been analysed and are presented in Figure 5.4. Whereas the water molecules on the surface do not appear to be strongly affected compared to the bulk molecules, there is definitely an increase in the diffusion coefficients of the surface ions. These can diffuse up to two orders of magnitude faster when compared to bulk ACC. An

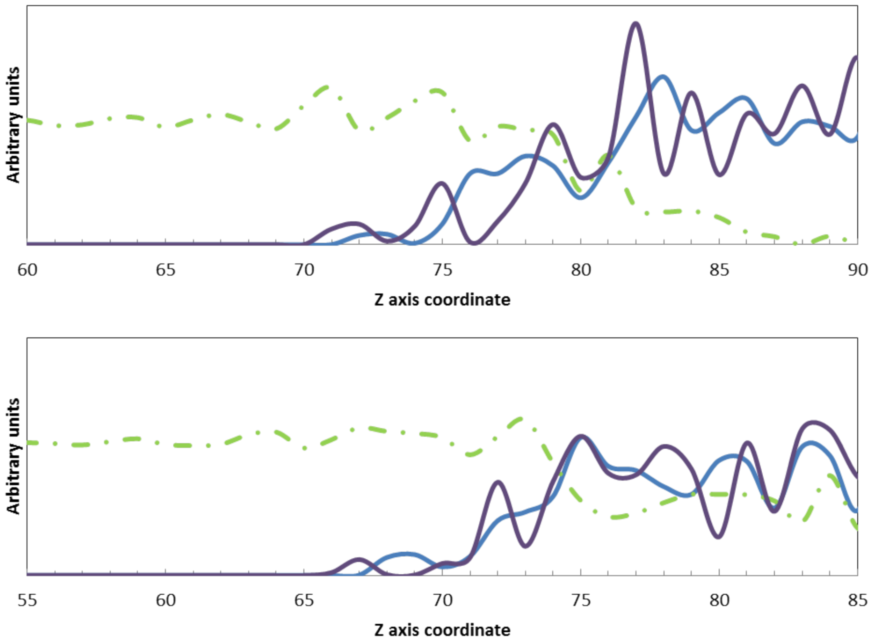


Figure 5.3: Density profiles of water (green dashed line), carbonate (purple) and calcium (blue) along the Z axis. The top image represents ACC_1 and the bottom ACC_2. In the graphs, the ACC slab occupy the central position, approximately from 80 until 120 Å.

increased mobility suggests that ions could restructure around incoming molecules and/or that formation of a crystalline nuclei is potentially more likely. Interestingly, there is not a large difference in the diffusion coefficient of the ions between ACC_1 and ACC_2, suggesting that the presence of water within the structure does not have an important effect on the dynamics of the system. The large peaks present in Figure 5.4 far away from the surface correspond to ions that diffused into the solution.

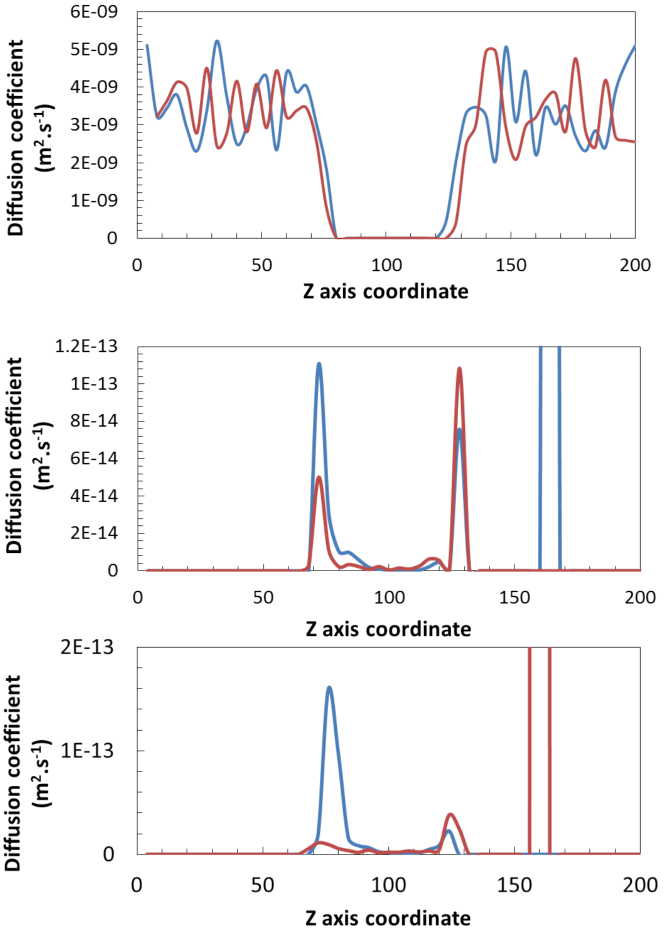


Figure 5.4: Diffusion profiles of the different elements of the system along the Z axis. From top to bottom: O_w, Ca^{2+} and CO_3^{2-}. The blue line depicts the result obtained from ACC_1 and the red line from ACC_2. Note that in the two bottom figures, the peak far from the surface is due to ions that detached from the ACC. Thus, there diffusion coefficient increased in magnitude.

5.3.2 Configurational binding energies of amino acids on ACC

The configurational binding energies obtained from both the hydrous and anhydrous surfaces are presented in Table 5.1. The amino acid with the most favourable binding energy on ACC_1, is arginine followed by glutamic acid, glycine and aspartic acid. On ACC_2 the results are slightly different. The most favourable amino acid is now glycine, glutamic acid is second followed by aspartic acid and then arginine.

Glutamic acid, glycine and arginine all show a clear preference for one of the surfaces. Both arginine and glutamic acid appear to bind more favourably on ACC_1 while glycine prefers ACC_2. By contrast aspartic acid binds with approximately the same configurational energy to both surfaces.

As shown experimentally, both glutamic acid and glycine are strong binders of the amorphous surface. These amino acids are often present together in proteins which are known to stabilize ACC [197]. Intriguingly, one shows a stronger configurational binding energy to the anhydrous surface, while the other one to the hydrous surface. This could allow a protein to bind on both surfaces. By contrast aspartic acid shows an average binding to both surfaces. The difference in binding energy between glutamic and aspartic acid must lie in the higher flexibility of the former. Glutamic acid has an additional carbon atom in the backbone of the side chain, which allows the molecule to find more favourable interactions with the surface.

Table 5.1: Configurational binding energies ($kJ.mol^{-1}$) of amino acids on ACC surfaces.

Molecule	ACC_1	ACC_2
Aspartic acid	-411.02 ± 1.16	-418.82 ± 1.12
Arginine	-901.12 ± 1.06	-342.29 ± 2.12
Glutamic acid	-705.78 ± 1.09	-496.77 ± 1.05
Glycine	-463.34 ± 0.91	-771.56 ± 1.11

5.3.3 Binding motifs of amino acids on ACC

Figures 5.5 and 5.6 illustrate the interactions between the amino acids and the two ACC surfaces. It is clear from both Figures that the amine and guanidine groups are the preferred binding motifs (see also the density profiles in Figures 5.7 and 5.8). The carboxylic groups only rarely interact with the surfaces. Even in the cases where the molecule spreads on the surface, like glutamic acid on ACC_2 (see Figure 5.8), there is a sizeable gap filled with water molecules between the carboxylic group and the surface. Interaction via the amine or guanidine groups, however, always shows hydrogen bonding between the H_n (amine or guanidine hydrogen) and the O_c. It is surprising that in general the carboxyl group has not been observed to interact with the surface, as this is usually believed to be highly important [195]. The reason could lie in a considerable energy barrier arising from the pronounced solvation shell of the calcium ion. Such a barrier was observed for aspartate by Raiteri *et al.* [200] and for acetate by Saharay and Kirkpatrick [176]. In the latter study, no energetic barrier was observed for the binding of an ion pair to guanidine. Additionally, in this study, the change in Gibb's free energy measured for acetate and guanidine binding

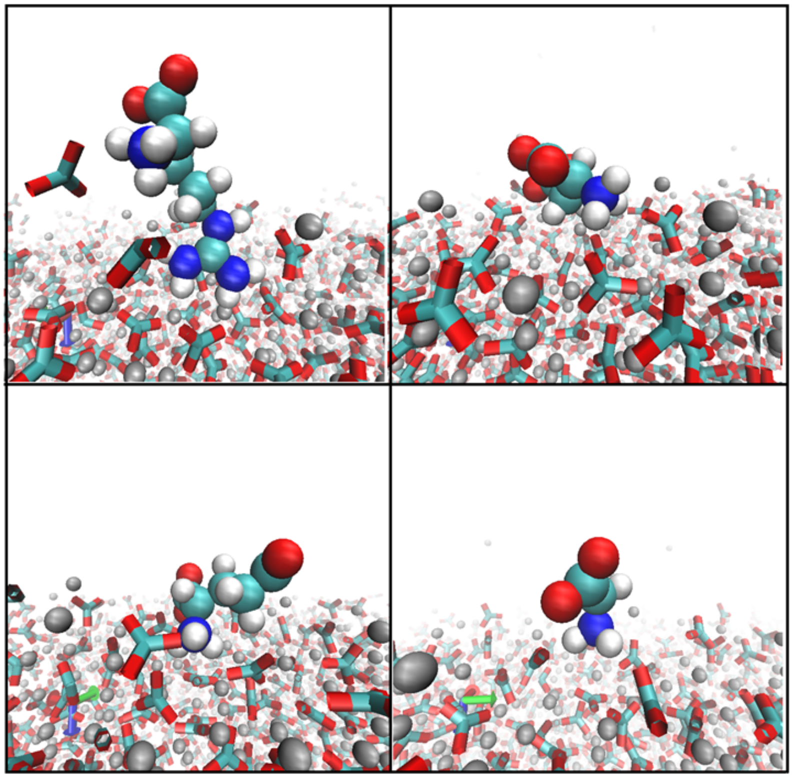


Figure 5.5: Amino acids on the surface of ACC_1. From left to right and top to bottom: arginine, glutamic acid, aspartic acid and glycine. Water was removed from the images for clarity.

to a calcium carbonate ion pair was approximately the same, 5 $kJ.mol^{-1}$ [176]. Interaction between guanidine and ACC should therefore be preferred from a kinetic point of view. It can be speculated, in light of the present observations, that a similar behaviour for the amine group is likely.

The density profiles presented in Figures 5.7 and 5.8 show that some of the molecules are slowly starting to be incorporated within the structure. Often the increase in density of the amines is closely followed by the carbonates. The surface, being more mobile than

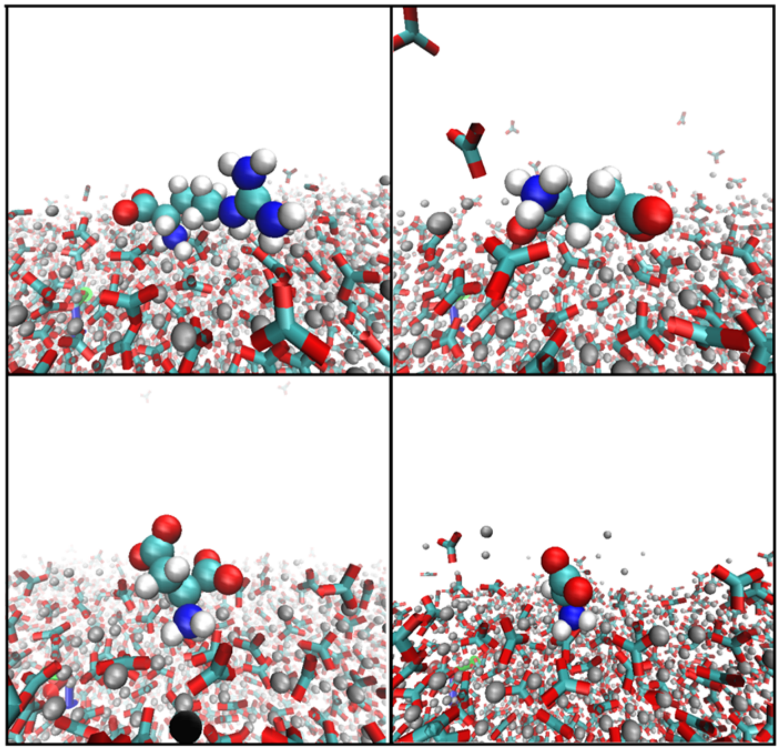


Figure 5.6: Amino acids on the surface of ACC_2. From left to right and top to bottom: arginine, glutamic acid, aspartic acid and glycine. Water was removed from the images for clarity.

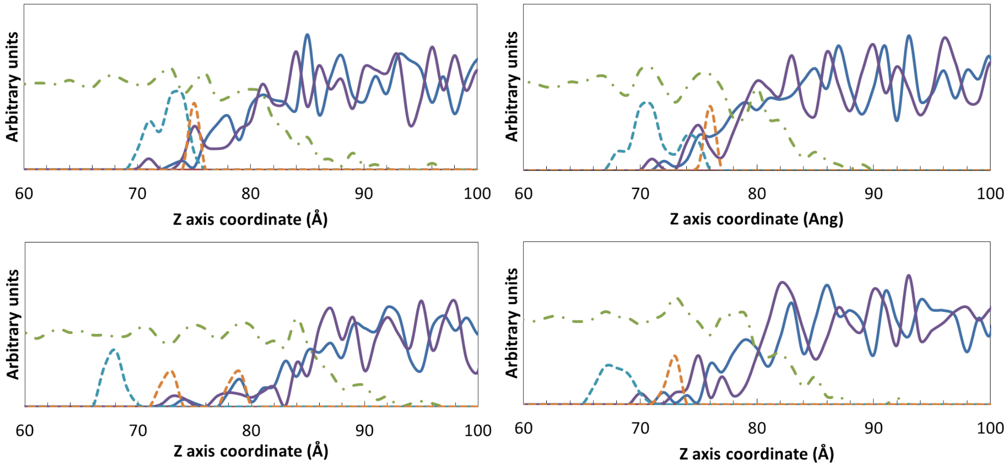


Figure 5.7: Density profiles of waters (green dashed line), carbonates (purple line), calciums (blue line), carboxyl oxygens (blue dashed line) and nitrogens (orange dashed line) obtained from ACC_1. From left to right and top to bottom: aspartic acid, glutamic acid, arginine and glycine.

the bulk, reacts to the presence of the amino acid by optimizing the position of the ions. This allows a maximum interaction between the O_c and the H_n of the amine or guanidine groups.

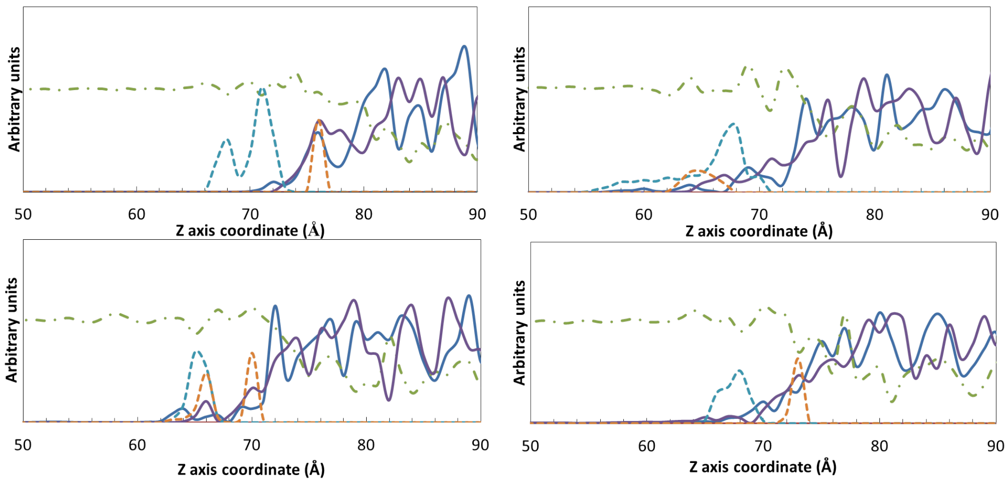


Figure 5.8: Density profiles of waters (green dashed line), carbonates (purple line), calciums (blue line), carboxyl oxygens (blue dashed line) and nitrogens (orange dashed line) obtained from ACC_2. From left to right and top to bottom: aspartic acid, glutamic acid, arginine and glycine.

5.3.4 Water structure at surface and effect of amino acids binding

As recognised by Dorvee and Veis, the residence time of water molecules in ions coordination shells is an important consideration for the kinetics of nucleation, growth and crystallization [209]. A decrease in the water residence time would indicate that the energy barrier present at the transition between coordination shells is smaller, and thus that interaction with other elements will be facilitated. Organic molecules are known to modify water structure in their surroundings [9], and thus they might also be able to affect its dynamics.

The residence time of water molecules on the surface of ACC_1 without amino acids ranged from 2 ps to 288 ps. The former value match the residence time calculated for water molecules within the coordination shell of the carbonate ion in bulk solution [209]. This must be a ion which is highly hydrated but still on the surface. The latter value is within the range found for water molecules in the coordination shell of calcium ions [209]. As shown in Figure 5.9, these water molecules are electrostatically interacting with the calcium ion while forming an hydrogen bond with the carbonate ion. The result is a stable configuration and the molecule will be "constrained" in the same position for a non negligible amount of time. The average residence time found for this surface was 19.8 ps, which is approximately 10 ps smaller than the residence time of water molecules on the surface of calcium carbonate crystalline nanoparticles, computed using a different force field [210, 211]. In addition it is also within the experimental values obtained using ultrafast infrared spectroscopy measurements. The interfacial water orientational relaxation time was measured to be 18 ± 3 ps close to the interface, while being 2.6 ps in bulk solution. Importantly, the chemical nature and charge of the surface were shown not to have an impact. Similar results were obtained for surfaces composed of sodium bis(2-ethylhexyl) sulfosuccinate (electrically charged), Igepal (neutral) and tetraethylene glycol dimethyl ether micelles [212].

The results obtained for the residence time of water molecules on ACC_1 when the amino acid was present are plotted on a 3D diagram shown in Figure 5.10. In general, there is a clear correlation between the residence time of the water molecules and the position of the amino acid on the surface. The presence of the amino acid leads to a decrease in the residence time of the water molecules and thus a local increase in the dynamics of the system. The root lies both in the effect of the amino acid on the structure of the surrounding

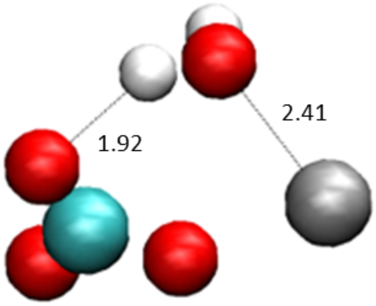


Figure 5.9: Example of a water molecule with a long residence time in the coordination shell of a O_c. The water is forming an hydrogen bond with the O_c and is within the coordination shell of a Ca^{2+}. In the image the oxygens are depicted in red, the carbon in blue, the hydrogens in white and the calcium is coloured in grey.

water molecules and in the way that it binds to the surface. As shown in Figures 5.7 and 5.8, the binding occurs mainly through hydrogen bonds between H_n and O_c. The latter have a lower residence time for water molecules in their coordination shell than calcium ions. As amino acids bind to, or lead to, regions of higher carbonate ion concentration, the residence time of the local water molecules will be lower in that region.

Hydrous ACC_2 shows a similar average residence time of the water molecules compared to ACC_1. The average here is 18.1 ps with a maximum value of 194 ps and a minimum value of 2 ps. So in general, water appears to behave similarly on both surfaces. This is sensible as the ACC_1 surface quickly incorporates water, and thus, must be chemically similar to ACC_2. The results, when the amino acids are present on the surface of ACC_2, are not as clear as for ACC_1. Qualitative trends can, however, be observed in Figure

5.11. In the case of aspartic acid, glutamic acid and glycine it seems that the residence times are slightly increased close to the molecule compared to other areas of the surface. In contrast, arginine still seems to show the same effect seen on the anhydrous surface; a decrease in the residence time of the water molecules. The reason behind the difference in the effect of amino acids on ACC_1 and ACC_2 is unclear. The number of oxygens from the water molecules or the carbonate ions within the coordination shell of the H_n does not show tangible differences between the two surfaces. The water molecules that are intrinsic to the surface must be leading to changes in the structure, which affect the 'free' water molecules and the binding of amino acids. Such an effect could be an additional reason for the precipitation of this hydrous form *in vivo*. Not only is this form of ACC more stable than the 'anhydrous' counterpart, but interactions with amino acids can increase the residence time of water molecules on the surface. This would further hinder transformation or dissolution towards the formation of a more stable polymorph.

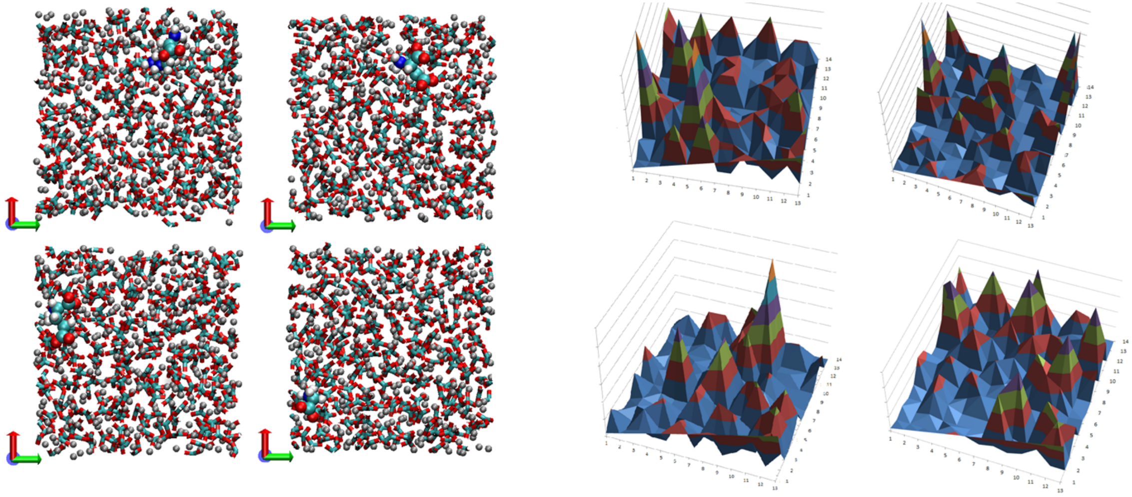


Figure 5.10: The left-hand four images show the position of the amino acids on a ACC_1 surface. The x axis is the red arrow while the green arrow is the y axis. The right-hand four images show the residence time of the water molecules on the surface. They are oriented in the same way as the surfaces. In these diagrams, the z axis shows the magnitude of the residence time. From left to right and top to bottom: arginine, aspartic acid, glutamic acid and glycine.

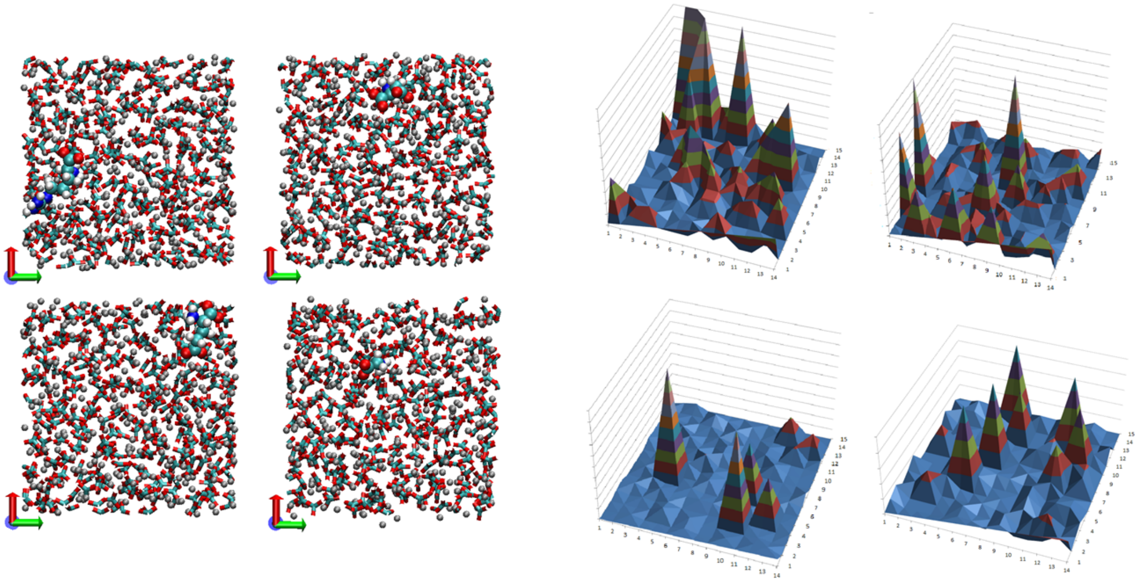


Figure 5.11: The left-hand four images show the position of the amino acids on a ACC_2 surface. The x axis is the red arrow while the green arrow is the y axis. The right-hand four images show the residence time of the water molecules on the surface. They are oriented in the same way as the surfaces. In these diagrams, the z axis shows the magnitude of the residence time. From left to right and top to bottom: arginine, aspartic acid, glutamic acid and glycine.

5.4 Discussion

The ACC surfaces are still poorly defined and experimental studies analysing this important interface are lacking. Here we show that the surface is diffuse and does not lead to a strong structuring of the water layers as seen in the crystalline material. Additionally, the results obtained are comparable to previous observations gathered using a different model for the calcium carbonate system [9]. This suggests that the water structure at the interface of ACC is only weakly dependent upon the chosen force field.

All amino acids analysed show a favourable configurational binding energy to the surface. This might be a property of the amorphous surface, where the ions can diffuse two orders of magnitude faster than in the bulk. The result is that the ionic framework can reorganize to maximize the favourable interactions with the incoming molecules. A number of pathways will thus be available for small molecules to interact with a surface. Additionally, previous calculations by Raiteri *et al.*, using nanoparticles of 36 formula units, showed that the barrier to attachment of aspartic acid and acetate was less than 2 $kJ.mol^{-1}$, which is below thermal energies at room temperature (about 2.5 $kJ.mol^{-1}$). Both molecules also showed a favourable change in Gibb's free energy upon binding [200]. Experimentally all of the above amino acids were found in proximity to the highly stable ACC of earthworms [202]. The most favourable binding energy for the anhydrous surface was observed for arginine, a basic amino acid. These are not usually considered important in biomineralization, as they are found sparingly in protein sequences involved in the process. However, this might be the reason behind their rare observation. As they can bind so tightly to the surface, they could potentially lock the protein and not allow for a subsequent processing, which might require another organic molecule. The study of Saharay and Kirkpatrick showed that the free energy released from the attachment of calcium carbonate to either guanidine or acetate was the same. The difference between the two molecules resided in the energy barrier at the transition state. The free energy barrier for the association with acetate was approximately twice as large as the thermal energy fluctuations at room temperature, while the one for guanidine was approximately equal. Thus, it is easier for positive charges to interact with the surfaces. A similar result is observed in the present simulations, where all of the amino acids bind via the amine group on both ACC_1 and ACC_2.

Analysing the residence time of the water molecules on the surface has led to some interesting results. The effect of the amino acids on the two different surfaces was distinct. On ACC_1 the amino acid decreased the average residence times of water molecules, while on ACC_2 in general they seemed to increase them. However, this result is important as it gives a clue to a potential mechanism of stabilisation and destabilisation of ACC that depends on its hydration level. Hydrous ACC could be precipitated by living organisms, because it can be easily further stabilised by molecules present in the solution. By contrast, transient ACC_1 can be further destabilised by interactions with organic molecules.

5.5 Conclusion

All amino acids appear to be able to bind to ACC. The fact that the water is weakly structured at the surface, suggests that the barrier on the free energy curve when passing through the hydration shell must be low. This transition barrier will be highly dependent upon the local ACC structure, and in some cases could be absent. Additionally, the penetration of water within the surface and the partial dissolution of ions leads to a dynamic surface. The diffusion coefficients are two orders of magnitude larger than those of the bulk and thus the surface's ions can cluster around the molecule to maximise favourable interactions.

In all the configurations sampled, binding occured mainly via the amine and guanidine groups. It is thus not possible to extrapolate the results of the acid amino acids to the full organic polymer. Within proteins, the amine group is part of the backbone where it is involved in the peptide bond. It is unlikely that it will be able to interact with the surface. Basic groups both show a strong interaction and could have a lower barrier on the free energy curve of binding. This raises important questions as to their role in protein attachment to surfaces. Positively charged groups in the protein could be essential in the initial binding of a protein to a surface. Anchored to the surface, the protein could then find the most favourable interactions, while ions in proximity would also diffuse towards optimized positions.

5.6 Future work

It is clear that the next important step is to measure the free energy change as a function of the distance between the amino acid and the surface. Even if the ensemble of potential pathways at the transition cannot be sampled, measurements in different regions could lead to an accurate estimate.

Analysing the effect of the amino acids on crystallization could give some interesting insights into polymorph selection. Such simulations could be carried using advanced sampling techniques such, as metadynamics. The result would be a free energy map along selected order parameters, that could show differences depending on the amino acid present. Additionally, the free energy map as a function of the Q_4 Steinhardt order parameter for nanoparticles in solution has not yet been obtained with the most recent calcium carbonate force field [87]. This should be able to capture more accurately the free energy difference between calcite and aragonite. The latter was not observed in previous free energy surfaces obtained using the Pavese potential [174].

Chapter 6

Conclusion

The initial aims of the present investigation were to gain a comprehensive understanding of the molecular details of the nucleation and growth of minerals. As the study was part of a wider investigation to understand biomineralization and the interaction between soft and hard matter, this work primarily focussed on two of the most widespread mineral systems: calcium phosphate and calcium carbonate.

There is still an open debate about the nucleation and growth of biominerals. Recent observation of the formation of negative complexes and amorphous precipitates is intriguing. The study of the thermodynamic stability of these species is important to investigate if they could be part of a classical nucleation pathway, non-classical pathway, or if they actually inhibit the subsequent nucleation. Another important step in biomineralization is the formation of amorphous precursors. In the case of $CaCO_3$ the structure of amorphous phases has yet to be fully understood, with crystallization mechanisms and the role of organic molecules on this process proving to be controversial.

6.1 Nucleation of biominerals

The results obtained from Molecular Dynamics simulations showed that negatively charged clusters, including the $[Ca(HPO_4)_3]^{4-}$ complex described by Habraken *et al.*, are (meta)stable compared to free ions in solution [56]. The most stable complex, however, from the results obtained in this investigation is $[Ca(HPO_4)_2]^{2-}$, consistent with the results obtained by Zhang using EDXA *et al* [136]. These complexes can then aggregate to form larger metastable complexes with a barrier to formation of approximately 10 $kJ.mol^{-1}$ from the monomer units. The fact that these aggregates occupy local minima rather than global ones is extremely important. Following classical nucleation theory, these aggregates could then lie on the nucleation pathway as they do not lead to a larger nucleation barrier. Instead, they would be a step closer towards reaching the nucleation transition state. This is in contrast to the calcium carbonate system. In this case the addition of ion pairs to "pre-nucleation" species was suggested to occur without a transition barrier and lead to a minima more stable compared to the monomers. Thus it would be even harder for this species to overcome the nucleation transition barrier, at least for mechanisms that are conceivable within classical nucleation theory.

The results obtained here show that the pathway envisaged experimentally for the calcium phosphate system is indeed feasible and additionally highlight the importance of solute molecules in the solution. Increasing the ionic strength in the simulations led to the formation of a metastable state for the aggregation of two $[Ca(HPO_4)_3]^{4-}$ complexes.

The presence of TRIS, instead, allowed the aggregation of three $[Ca(HPO_4)_3]^{4-}$ complexes for prolonged periods of time. The relative free energy change of the latter reaction has, however, not been measured. TRIS could have an important impact in biomimetic studies, and its effect has rarely been studied. More generally, from the simulations, it seems that most additives can impact the potential pathways to nucleation.

It is apparent from both the simulations and the new experimental results that speciation calculations of calcium phosphate solutions need to be repeated. If ionic complexes do exist in solution they need to be taken into consideration, which could lead to considerable changes in the relative equilibrium constants of ion pairing. The effect of TRIS on the process should always be taken into account. This study shows that TRIS assists the formation of large aggregates. Hydrogen bonding between hydrogen phosphate and TRIS molecule, appeared extremely favourable in our simulations, and measuring the free energy of association will be important to understand the potential population of hydrogen phosphate bond to TRIS in experiments.

6.2 Structure of the hydrous amorphous precursor to calcium carbonate polymorphs

ACC has recently gained a large amount of interest due to its presence as a precursor to crystalline phases in most mineralizing living organisms. Its properties are important for the subsequent determination of the crystalline polymorph, the morphology of the final crystal and the growth rate of the crystalline phase. While efforts to study the structure of

ACC both experimentally and using computer simulations have produced a large quantity of data, there are still questions about both the structure and the role of water molecules within hydrous ACC. As ACC has been observed with different short range structures, there must be a number of local minima in the free energy landscape of this phase. Thus, it is important to sample the configurational space more comprehensively to ensure that the structure being studied is statistically significant. The simulations investigated in Chapter 4 used four significantly different protocols to obtain ACC models. All the models led to structure which were in good agreement with experimental observations, including the PDF and the coordination shells of the different ions.

While three of the generated ACC models showed very similar properties, one, ACC_{Ik} was very different. Percolating channels were observed in this structure. There, the water molecules showed diffusion coefficients four orders of magnitude greater than in the bulk. Such channels could be pathways for water expulsion. This structure, however, was higher in energy compared to the other sampled configurations and thus will have a low probability in the ensemble of configurations. The formation of microscopic channels would be possible via kinetic processes, which would mean that water could be trapped in mesoscopic channels during precipitation/aggregation. Overall, an homogeneous mixture of water within ACC is more probable. It is likely that the dehydration of ACC happens through an alternative mechanism which would either involve dissolution and reprecipitation or, a collective reorganization of the ions.

It is interesting to note that the PDF of ACC obtained for all sampled structures was identical. As shown in Chapter 5, even if water is completely removed from the ACC the PDF is not affected. A large number of structurally different configurations can thus be obtained with the same elements, all mapping onto the same PDF. Additionally, the observation suggests that the short range order is only determined via the coordination shell of Ca^+ ions to oxygens which always fall between 7 and 8 both for hydrated and anhydrous ACC.

All four ACC models fell within the experimentally obtained observations. Experiments, however, show a much wider range in the order parameters examined. The model obtained in this study can now be used as a benchmark to understand the effect of additional solutes both on the water and ionic structure. This could shed light on the mechanism of crystallization and polymorph selection.

6.3 Effect of organic molecules on growth and crystallization

Organic molecules are part of a living organism's toolbox to achieve a myriad of processes; from the transport of oxygen around the body to the control of muscle contraction. In the case of biomineralization these are essential in modifying the free energy landscape and thus play a pivotal role both in the nucleation and growth of mineral phases. The investigation in Chapter 5, focussed on the binding of four amino acids to both hydrous and anhydrous ACC.

All amino acids showed a favourable binding configurational energy to ACC. Additionally, both the water molecules and the ions can diffuse at the interface, leading to many potential pathways for the binding process, which might not present a transition barrier. The study of Raiteri *et al.* on nanoparticles composed of 36 formula units of $CaCO_3$ showed a barrier which was below room temperature energy fluctuations [200]. As the number of pathways to the adsorption of amino acids on such small nanoparticles is small compared to those present at an ACC surface it is probable that lower energy transition barriers would be present.

The amino acid with the strongest binding energy was arginine on anhydrous ACC. The fact that arginine bound through its positively charged guanidinium group was most surprising. However, all the amino acids were observed to bind via their amine group, so generally, it seems, that binding via hydrogen bonds is the favoured mode of interaction. This is also supported via calculations using a completely different force field for the organic molecules. The result is that such a mode of interaction could be a general binding motif in living organisms. The lack of basic functional groups within biomineralizing proteins could arise from a need to avoid tight binding to the mineral surfaces. Proteins and enzymes within living organisms, are known to work as catalysts cycle and Freeman *et al.* suggested that ovocleidin-17 can catalyse the crystallization of ACC nanoparticles [198].

The amino acids tested were all observed to affect the dynamics of the water molecules present on the surface of ACC. It was surprising to see the difference between the hydrous and anhydrous surface. While both surfaces appeared structurally similar, on anhydrous ACC, which is usually considered to be transient in solution, the amino acids increased the dynamics of the water molecules. This could increase the rates of either dissolution or precipitation depending on the supersaturation of the solution. Additionally, it would increase the chances for the formation of nuclei at the surface potentially starting crystallization. On the hydrous structure, which is considered to be more stable, three out of four amino acids led to a decrease in the dynamics of the water molecules at the surface. This would further stabilize the hydrous ACC.

Living organisms are able to alter the stability the stable and transient phases of ACC. How stable hydrous ACC can be destabilized needs to be explored in more detail. The simulations here clearly highlight the potential effect of organic molecules on biomineralization and showed that the binding via a positive charge is preferred on ACC.

Appendices

Appendix A

Force fields

All parameters are given in $kJ.mol^{-1}$ unless specified.

Note that in the following force fields parameters the Lennard-Jones potential is defined as:

$$U_{lj} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \tag{A.1}$$

where σ represent the distance at which the potential energy is equal to zero and ϵ is the depth of the minimum in the function. Additionally, the Lennard-Jones potential can be defined as:

$$U_{lj} = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}} \tag{A.2}$$

where $A = 4\epsilon\sigma^{12}$ and $B = 4\epsilon\sigma^{6}$. This form of the Lennard-Jones potential will be defined as $12 - 6$ in the following force fields definitions.

A.1 Water force field

Table A.1: Intramolecular parameters for the SPC/F water model developed by Wu, Tepper and Voth(2006) [17]. The charges of the O_W and the H_W atoms were -0.82 and 0.41 e, respectively.

Two body term		k	r_0(Å)
O_W-H_W	Harmonic	4431.6433	1.012
Three body term		k	θ
H_W-O_W-H_W	Harmonic	314.6785	113.24

Table A.2: Intermolecular parameters for the SPC/F water model developed by Wu, Tepper and Voth(2006) [17].

Lennard-Jones	ϵ	σ
O_W-O_W	0.6503	3.1655

Table A.3: Intramolecular parameters for the TIP3P water model developed by Jorgensen *et al* [18].The charges of the O_W and the H_W atoms were -0.834 and 0.417 e, respectively.

Two body term		k	r_0(Å)
O_W-H_W	Harmonic	3768.12	0.9572
Three body term		k	θ
H_W-O_W-H_W	Harmonic	460.52	104.52

Table A.4: Intermolecular parameters for the TIP3P water model developed by Jorgensen *et al* [18].

Lennard-Jones	ϵ	σ
O_W-O_W	0.6368	3.1506

A.2 Calcium phosphate force field

A.2.1 First parameters

Table A.5: Partial charges for the calcium phosphate force field model [19]

Atom	Partial charge (e)
Ca	2.0000
P	0.9700
O_P	-0.8900
O_{PH}	-0.6550
H_P	0.3550

Table A.6: Intramolecular parameters for the hydrogen phosphate ion [19].

Two body term		k	r_0(Å)	
P-O_P	Harmonic	4162.5606	1.5200	
P-O_{PH}	Harmonic	2338.27318	1.6000	
O_{PH}-H_P	Harmonic	4163.6606	0.9800	
Three body term		**k**	**θ**	
O_P-P-O_P	Harmonic	1145.78194	109.50	
O_{PH}-P-O_P	Harmonic	637.7688	109.50	
O_{PH}-H_P-P	Harmonic	280.3427	111.40	
Three body term		**A**	**d_{ab}**	**d_{ac}**
O_P-C-O_P	Compass stretch-stretch	156.2991	1.5200	1.5200
O_{PH}-C-O_P	Compass stretch-stretch	280.7866	1.6000	1.5200
Four body term		**A**	**δ**	**m**
H_P-O_{PH}-P-O_P	cosine	0.4079	0.0000	3.0000

Table A.7: Intermolecular parameters for the calcium phosphate mineral system [19]. The water potential used is the SPC/F water model developped by Wu, Tepper and Voth(2006) [17].

Lennard Jones	ϵ	σ	
Ca-O_W	0.09166	3.3500	
12-6	A	B	
P-O_W	305056.2153	0.2715	
H_P-O_P	5210.2926	0.0	
H_P-O_{PH}	5210.2926	0.0	
Buckingham	A	rho	C
O_P-O_W	1209410.7190	0.1550	0.0000
O_{PH}-O_W	1209410.7190	0.1550	0.0000
O_{PH}-H_W	38372.8401	0.2224	0.0000
O_P-H_W	74289.3168	0.2224	0.0000
H_P-O_W	6159742.8969	0.1989	0.0000
O_P-O_P	1209410.7190	0.2020	0.0000
O_P-O_{PH}	1209410.7190	0.2020	0.0000
O_{PH}-O_{PH}	1209410.7190	0.2020	0.0000
O_P-Ca	249985.7865	0.2763	0.0000
O_{PH}-Ca	194023.3845	0.2763	0.0000
P-Ca	205804.4929	0.3428	0.0000

A.2.2 Second parameters

Table A.8: Partial charges for the calcium phosphate force field model [19]

Atom	Partial charge (e)
Ca	2.0000
P	0.9700
O_P	-0.8900
O_{PH}	-0.6550
H_P	0.3550

Table A.9: Intramolecular parameters for the hydrogen phosphate ion [19].

Two body term		k	r_0(Å)	
P-O_P	Harmonic	4162.5606	1.5200	
P-O_{PH}	Harmonic	2338.27318	1.6000	
O_{PH}-H_P	Harmonic	4163.6606	0.9800	
Three body term		**k**	**θ**	
O_P-P-O_P	Harmonic	1145.78194	109.50	
O_{PH}-P-O_P	Harmonic	637.7688	109.50	
O_{PH}-H_P-P	Harmonic	280.3427	111.40	
Three body term		**A**	**d_{ab}**	**d_{ac}**
O_P-C-O_P	Compass stretch-stretch	156.2991	1.5200	1.5200
O_{PH}-C-O_P	Compass stretch-stretch	280.7866	1.6000	1.5200
Four body term		**A**	**δ**	**m**
H_P-O_{PH}-P-O_P	cosine	0.4079	0.0000	3.0000

Table A.10: Intermolecular parameters for the calcium phosphate mineral system [19]. The water potential used is the SPC/F water model developped by Wu, Tepper and Voth(2006) [17].

Lennard Jones	ϵ	σ	
Ca-O_W	0.09166	3.3500	
12-6	A	B	
P-O_W	305056.2153	0.2715	
H_P-O_P	5210.2926	0.0	
H_P-O_{PH}	5210.2926	0.0	
O_W-H_P	723.65175	0.0	
H_W-O_{PH}	2701.6332	0.0	
H_W-O_P	1350.8166	0.0	
Buckingham	A	rho	C
O_P-O_W	732335.571	0.2301	0.0000
O_{PH}-O_W	732335.571	0.2151	0.0000
H_P-O_W	6159742.8969	0.1989	0.0000
O_P-O_P	1209410.7190	0.2020	0.0000
O_P-O_{PH}	1209410.7190	0.2020	0.0000
O_{PH}-O_{PH}	1209410.7190	0.2020	0.0000
O_P-Ca	249985.7865	0.2763	0.0000
O_{PH}-Ca	194023.3845	0.2763	0.0000
P-Ca	205804.4929	0.3428	0.0000

A.3 Calcium carbonate force field

Table A.11: Partial charges for the calcium phosphate force field model [19]

Atom	Partial charge (e)
Ca	2.0000
C	1.1233
O_C	-1.0411

Table A.12: Intramolecular parameters for the carbonate ion [20].

Two body term		k				
			r_0(Å)			
C-O_C	Harmonic	3941.4223	1.3042			
Three body term		k	θ			
O_C-C-O_C	Harmonic	1276.9076	120.0			
Three body term		A	B	C	θ_0	$d_{ab} = d_{ac}$
O_C-C-O_C	Compass all terms	1236.7691	147.9327	147.9327	120.0	1.3042
Inversion angle		k	Φ			
C-O_C-O_C-O_C	Calcite	1316.7567	360.0			

Table A.13: Intermolecular parameters for the calcium carbonate system [20]. The water potential used is the SPC/F water model developped by Wu, Tepper and Voth(2006) [17].

Lennard Jones	ϵ	σ	
O_W-O_W	0.6503	3.1655	
Ca-O_W	0.09166	3.3500	
Buckingham	**A**	**rho**	**C**
O_C-O_W	1209410.7190	0.2020	1166.5483
O_C-H_W	38237.7585	0.2170	0.0000
O_C-O_C	6159742.8969	0.1989	2691.8890
Ca-O_C	305056.2153	0.2715	0.0000

A.4 Organic force fields

As AMBER is a standard package, in this appendice, we only include the pair-pair intermolecular potential for the organic force fields and the intermolecular interactions between the organic and mineral force field as obtained using the Schroeder method [21]. The method used to determine the intramolecular parameter is detailed in Section 5.2.2.

A.4.1 TRIS

Table A.14: Partial charges for the TRIS molecule. Obtained using the bcc method which is part of the AMBER package.

Atom	Partial charge (e)
c_3	0.3090
c_3	-0.1401
c_3	-0.1763
c_3	-0.1597
n_4	-0.6745
O_H	-0.5402
O_H	-0.5195
O_H	-0.5020
h_1	0.1412
h_1	0.2089
h_1	0.1825
h_1	0.1479
h_1	0.1809
h_1	0.1367
h_n	0.4009
h_n	0.4215
h_n	0.4325
H_O	0.3876
H_O	0.3886
H_O	0.3740

Table A.15: Intermolecular parameters for the TRIS molecule as obtained from AMBER. Only the pair-pair potentials are given. All the other intermolecular parameters can be obtained using the Lorentz-Berthelot rules.

12-6	A	B
c_3-c_3	4367168.3070	2828.1995
n_4-n_4	3953566.9079	3354.4612
O_H-O_H	2435893.7592	2929.2661
h_1-h_1	13647.6962	59.8959
h_n-h_n	0.5861	0.3925
n_4-n_4	3953566.9079	3354.4612

Table A.16: Intermolecular parameters between the TRIS molecule and the calcium ions that are calculated using the Schroeder rules as explained by Freeman *et al.* [21]. The pairs not present in the following tables may be obtained via the standard Lorentz-Berthelot rules.

Buckingham	A	ρ	C
Ca-n_3	19598.4038	0.4030	0.0000
Ca-O_H	149480.3087	0.2763	0.0000

A.4.2 Amino acids

Arginine

Table A.17: Partial charges for the arginine amino acid. Obtained using the bcc method which is part of the AMBER package.

Atom	Partial charge (e)
n_1	-0.8276
c_3	-0.0285
c_z	0.8976
o	-0.7903
c_3	-0.0914
c_3	-0.1024
c_3	0.0173
n_2	-0.3881
c_z	0.5113
n_3	-0.4807
n_3	-0.4807
o	-0.7903
h_n	0.4578
h_n	0.4578
h_n	0.4578
h_x	0.1157
h_c	0.0612
h_c	0.0612
h_c	0.0842
h_c	0.0842
h_1	0.0732
h_1	0.0732
h_n	0.3337
h_n	0.3232
h_n	0.3232
h_n	0.3232
h_n	0.3232

Table A.18: Intermolecular parameters for arginine as obtained from AMBER. Only the pair-pair potentials are given. All the other intermolecular parameters can be obtained using the Lorentz-Berthelot rules.

12-6	A	B
c_3-c_3	4367168.3070	2828.1995
c_z-c_z	3433057.3545	2223.2647
n_1-n_1	3953566.9079	3354.4612
o-o	1590466.5073	2364.7244
h_1-h_1	13647.6962	59.8959
h_n-h_n	0.5861	0.3925
h_x-h_x	844.9948	14.9041
h_c-h_c	31468.3113	90.9496
n_3-n_3	3953566.9079	3354.4612

Table A.19: Intermolecular parameters between arginine and the calcium ions, calculated using the Schroeder rules as explained by Freeman *et al.* [21]. The pairs not present in the following tables may be obtained via the standard Lorentz-Berthelot rules.

Buckingham	A	ρ	C
Ca-n_1	539548.0390	0.2530	0.0000
Ca-n_2	238220.1671	0.2530	0.0000
Ca-n_3	290722.8985	0.2530	0.0000
Ca-o	238647.5256	0.2715	0.0000

Aspartic acid

Table A.20: Partial charges for aspartic acid. Obtained using the bcc method which is part of the AMBER package.

Atom	Partial charge (e)
n4	-0.8656
c3	-0.0075
c	0.9116
o1	-0.8083
c3	-0.2014
c	0.8976
o2	-0.8333
o2	-0.8333
o1	-0.8083
hn	0.4455
hn	0.4455
hn	0.4455
hx	0.0737
hc	0.0692
hc	0.0692

Table A.21: Intermolecular parameters for aspartic acid as obtained from AMBER. Only the pair-pair potentials are given. All the other intermolecular parameters can be obtained using the Lorentz-Berthelot rules.

12-6	A	B
c_3-c_3	4367168.3070	2828.1995
c-c	3433057.3545	2223.2647
n_4-n_4	3953566.9079	3354.4612
o_1-o_1	1590466.5073	2364.7244
o_2-o_2	1590466.5073	2364.7244
h_n-h_n	0.5861	0.3925
h_x-h_x	844.9948	14.9041
h_c-h_c	31468.3113	90.9496

Table A.22: Intermolecular parameters between aspartic acid and the calcium ions, calculated using the Schroeder rules as explained by Freeman *et al.* [21]. The pairs not present in the following tables may be obtained via the standard Lorentz-Berthelot rules.

Buckingham	A	ρ	C
Ca-n_4	508950.5448	0.2530	0.0000
Ca-o_1	243413.7445	0.2715	0.0000
Ca-o_2	250033.4931	0.2715	0.0000

Glutamic acid

Table A.23: Partial charges for glutamic acid. Obtained using the bcc method which is part of the AMBER package.

Atom	Partial charge (e)
n4	-0.8446
c3	-0.0225
c3	-0.0684
c3	-0.2024
c	0.8786
o1	-0.8583
o1	-0.8583
c	0.9176
o2	-0.8043
o2	-0.8043
hn	0.4598
hn	0.4598
hn	0.4598
hx	0.0787
hc	0.0652
hc	0.0652
hc	0.0397
hc	0.0397

Table A.24: Intermolecular parameters for glutamic acid as obtained from AMBER. Only the pair-pair potentials are given. All the other intermolecular parameters can be obtained using the Lorentz-Berthelot rules.

12-6	A	B
c_3-c_3	4367168.3070	2828.1995
c-c	3433057.3545	2223.2647
n_4-n_4	3953566.9079	3354.4612
o_1-o_1	1590466.5073	2364.7244
o_2-o_2	1590466.5073	2364.7244
h_n-h_n	0.5861	0.3925
h_x-h_x	844.9948	14.9041
h_c-h_c	31468.3113	90.9496

Table A.25: Intermolecular parameters between glutamic acid and the calcium ions, calculated using the Schroeder rules as explained by Freeman *et al.* [21]. The pairs not present in the following tables may be obtained via the standard Lorentz-Berthelot rules.

Buckingham	A	ρ	C
Ca-n_4	508950.5448	0.2530	0.0000
Ca-o_1	252284.2075	0.2715	0.0000
Ca-o_2	250033.4931	0.2715	0.0000

Glycine

Table A.26: Partial charges for glycine acid. Obtained using the bcc method which is part of the AMBER package.

Atom	Partial charge (e)
n4	-0.8356
c3	-0.0952
c	0.9326
o	-0.7553
o	-0.7553
hn	0.4438
hn	0.4438
hn	0.4438
hx	0.0887
hx	0.0887

Table A.27: Intermolecular parameters for glycine as obtained from AMBER. Only the pair-pair potentials are given. All the other intermolecular parameters can be obtained using the Lorentz-Berthelot rules.

12-6	A	B
c_3-c_3	4112161.9679	2663.0562
c-c	3433057.3545	2223.2647
n_4-n_4	3953566.9079	3354.4612
o-o	1590466.5073	2364.7244
h_n-h_n	0.5861	0.3925
h_x-h_x	844.9948	14.9041

Table A.28: Intermolecular parameters between glycine and the calcium ions, calculated using the Schroeder rules as explained by Freeman *et al.* [21]. The pairs not present in the following tables may be obtained via the standard Lorentz-Berthelot rules.

Buckingham	A	ρ	C
Ca-n_4	491949.8503	0.2530	0.0000
Ca-o	229379.8777	0.2715	0.0000

Appendix B

Calcium phosphate force field fitting

The data that will be discussed in this section has mostly been obtained by Gale *et al.* at Curtin University [19]. Gale *et al.* have provided us access to the unpublished version of the force field and the methods used to obtain the parameters [19].

Many force fields have been developed for classical molecular dynamics simulations of hydroxyapatite and other calcium phosphate mineral phases. Unfortunately, none of these models included a force field for the HPO_4^{2-} ion. In a solution at pH 7.4, this ion is the main species along with the $H_2PO_4^{1-}$ ion. Thus, it is essential to include it in all studies investigating mineral nucleation and growth. Recently, Raiteri and Gale [87], used experimentally obtained thermodynamic values to obtain a force field for the calcium carbonate system [87]. This force field leads to free energy differences between polymorph comparable to experiments, and to accurate solvation energies for the ions. Additionally, the simulated binding energies of the ions and the structural parameters of the hydrous and anhydrous polymorphs were well reproduced [87].

The force field for the calcium phosphate system was obtained following a similar philosphy. The fitting was done using different thermodynamic parameters obtained either from experimental results or *ab initio* calculations. In the following sections the fitting procedure is briefly described. Additionally, structural and mechanical parameters of multiple calcium phosphate phases obtained with the force field will be compared to experimental values.

B.1 Fitting method

A few constraints were placed on the fitting method. The previously obtained calcium carbonate force field and the calcium phosphate force field have to be consistent to allow subsequent investigations involving both anions. Thus the Ca^{2+} ion and its intermolecular parameter with water were kept the same as suggested by Raiteri *et al* [87]. This implies that the water model (SPC/Fw) must also be the same. For the calcium carbonate system different water models have beens used, including rigid ones, without affecting the reproducibility of key quantities [203] suggesting this should not directly affect the reliability of the results. The last constraint was the overall charges of the anion and cation. They must be equal to the formal ionic charges to ensure that the force field is transferable between liquid and solid phases while mantaining charge balance.

B.1.1 Intramolecular parameters for the hydrogen phosphate anion

An initial estimate of the intramolecular parameters was obtained by matching to *ab initio* calculations of ions vacuum simulations at the M06/6-311++G* level of theory. Then the equilibrium geometry was obtained with the intramolecular potential was refined against values calculated at the M06/6-31+G** level of theory using the SM8 solvent model simulated with the dielectric constant of water. The final values obtained are shown in Section A.2.

B.1.2 Intermolecular parameters for the hydrogen phosphate anion

To obtain the parameters between the hydrogen phosphate anion and the water molecules a two-step approach was taken. Initially, *ab initio* calculations at the M06/6-311++G** level of theory were used to obtain binding configurations of the water molecules to both the $(P-)O^-$ and $(P-)O-H$ groups. Then, the Lennard-Jones interactions between the water oxygens and the phosphate oxygens were refined to fit the free energy of solvation. It is important to note that previous investigations in the literature for the solvation energy of the HPO_4^{2-} anion are scarce. George *et al.* measured an enthalpy change of -1255 $kJ.mol^{-1}$ using experimental heats of solution [213] and Florian and Warshel obtained a solvation free energy of -1025 $kJ.mol^{-1}$ determined by *ab initio* calculation using the iterative Langevin dipoles solvation model [214]. A Doctoral thesis by Moser calculating the solvation free energy at the B3LYP/6-31G level of theory suggests that the best value is -1070 $kJ.mol^{-1}$. However, Gale *et al.* at the B3LYP/6-31+G level of theory using an explicit solvation model found that -1140 $kJ.mol^{-1}$ for their fitting procedure.

Hydrogen phosphate - calcium interactions were fitted using lattice dynamics calculations. Previous experimental results including crystallographic data, bulk moduli and elastic constants were all used to obtain an accurate intermolecula interaction.

All the above intermolecular interactions are presented in Section A.2.

B.2 Results

B.2.1 Ionic solvation energies

As mentioned above, for the Ca^{2+} ion, the same intermolecular interactions with the water molecules as per the ones obtained for the calcium carbonate system were used [19, 87]. Using the same parameters for the Ca^{2+} ion allow for an easier integration of the carbonate and phosphate force fields.

Two different intermolecular parameters between the water and the hydrogen phosphate ion were tested. The one discussed in Chapter 3, $ff1$, and $ff2$, the new force field obtained by refining the water structure surrounding the anion.

$ff1$ matched well to previous RDF data calculated by De Leeuw *et al.* using density functional theory and the number of water molecules around the hydrogen phosphate anion matched nuclear magnetic resonance results obtained by Mason *et al* [215]. Additionally, the solvation free energy was -1157 $kJ.mol^{-1}$. This value is only slightly more favourable than the best estimate by Gale *et al.* of -1140 $kJ.mol^{-1}$.

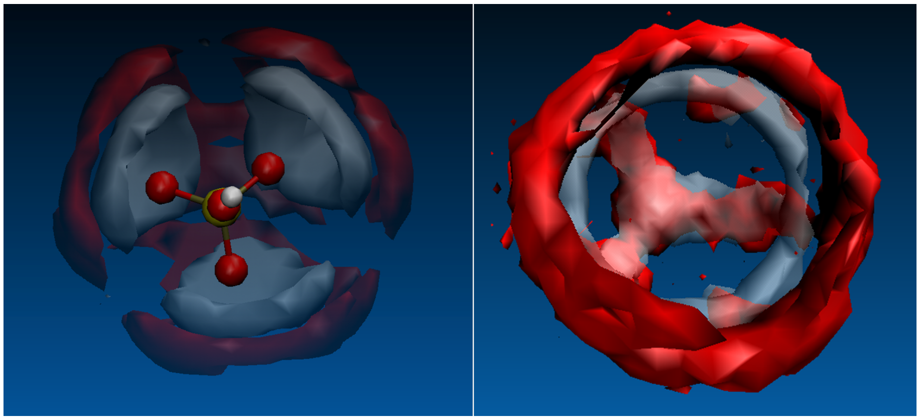


Figure B.1: Water density around the $ff2$ (on the left) and $ff1$ (on the right) force fields. The red surface depicts the average volume occupied by the water oxygens and the white surface is for the water hydrogens. In the right image the anion occupies the same position as in the left image and was removed for clarity.

However, additional calculations investigating the water density around the hydrogen phosphate anion using *ab initio* calculations showed a considerably different average distribution of the water molecules. In Figure B.1 the newly fitted intermolecular potential which match the *ab initio* water density, $ff2$, is compared to $ff1$. The latter leads to a configuration where the hydrogen of the water molecules sits between the phosphate oxygens, forming an hydrogen bond with both. This leads to an over structuring of the water molecules and an unrealistic hydrogen bond angle. In $ff2$, the water molecules interacts mainly with a single phosphate oxygen and the average density shows reasonable fluctuations surrounding the phosphate oxygens.

$ff2$ carries limitations as well. While it is able to capture the water structure accurately, it failed to obtain a solvation free energy within experimental error. The value obtained with this force field is approximately -1340 $kJ.mol^{-1}$.

B.3 Ability of force field to model calcium phosphate solid phases

B.3.1 Tricalcium phosphate

Table B.1: Experimental and calculated structural and mechanical properties for α-tricalcium phosphate. The experimental structure is taken from the work of Mathew *et al* [22].

Property	Experiment	This work	Hauptmans' force field [145]	Pedones' force field [146]	DFT calculations by Liang [216]
a (Å)	12.887	12.890	13.001	12.727	12.861
b (Å)	27.780	27.730	27.495	27.529	27.225
c (Å)	15.219	15.372	15.268	15.304	15.189
β (°)	126.20	126.81	125.31	126.86	126.20
K (GPa)	-	70.8	125.6	82.3	75.9
G (GPa)	-	33.9	60.4	36.8	40.1
C_{11} (GPa)	-	97.8	211.9	121.7	127.4
C_{22} (GPa)	-	149.0	202.5	156.1	135.3
C_{33} (GPa)	-	113.5	208.8	120.7	123.0
C_{44} (GPa)	-	36.4	60.7	34.2	40.2
C_{55} (GPa)	-	42.0	53.0	40.9	42.7
C_{66} (GPa)	-	27.7	71.1	39.4	39.9

B.3.2 Hydroxyapatite

Table B.2: Experimental and calculated structural properties for hydroxyapatite. The experimental structure is taken from the work of Catti *et al* [23].

Property	Experiment	This work	Error (%)
a (Å)	9.4166	9.3440	-0.77
b (Å)	9.4166	9.3440	-0.77
c (Å)	6.8748	6.9640	1.30
α (°)	90.00	90.00	0.00
β (°)	90.00	90.00	0.00
γ (°)	120.00	120.00	0.00

B.3.3 Monetite

Table B.3: Experimental and calculated structural properties for monetite. The experimental structure is taken from the work of Catti *et al* [23].

Property	Experiment	This work	Error (%)
a (Å)	6.910	6.981	0.94
b (Å)	6.627	6.661	0.94
c (Å)	6.998	6.905	-0.59
α (°)	96.34	93.24	-3.06
β (°)	103.82	104.03	0.20
γ (°)	88.33	90.10	1.99

B.4 Conclusion

The current force field shows a good fit to both the mechanical and structual parameters of a large number of mineral structures composed of calcium and hydrogen phosphate ions. The two intermolecular parametrization of the interaction between the hydrogen phosphate anion and water show some discrepancy with either structural or thermodynamic data. Work is still needed to ensure that both this properties are accurately captured to be able to confidently simulate the solution speciation.

Bibliography

[1] J. Y. Rho, L. Kuhn-Spearing, and P. Zioupos, “Mechanical properties and the hierarchical structure of bone,” *Med. Eng. Phys.*, vol. 20, no. 2, pp. 92–102, 1998.

[2] A. M. Bano, , P. M. Rodger, and D. Quigley, “New insight into the stability of $CaCO_3$ surfaces and nanoparticles via molecular simulation,” *Langmuir*, vol. 30, no. 25, pp. 7513–7521, 2014.

[3] D. Gebauer, A. Voelkel, and H. Coelfen, “Stable prenucleation calcium carbonate clusters,” *Science*, vol. 322, no. 5909, pp. 1819–1822, 2008.

[4] A. L. Goodwin, F. M. Michel, B. L. Phillips, D. A. Keen, M. T. Dove, and R. J. Reeder, “Nanoporous structure and medium-range order in synthetic amorphous calcium carbonate,” *Chemistry of Materials*, vol. 22, no. 10, pp. 3197–3205, 2010.

[5] A. V. Radha, A. Fernandez-Martinez, Y. Hu, Y. Jun, G. A. Waychunas, and A. Navrotsky, “Energetic and structural studies of amorphous Ca_1-xMgxCO_3.nH_2O (0 x 1),” *Geochimica Et Cosmochimica Acta*, vol. 90, pp. 83–95, 2012.

[6] Y. U. T. Gong, C. E. Killian, C. I. Olson, N. P. Appathurai, A. L. Amasino, M. C. Martin, L. J. Holt, F. H. Wilt, and P. U. P. A. Gilbert, "Phase transitions in biogenic amorphous calcium carbonate," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 16, pp. 6088–6093, 2012.

[7] Y. Oaki and H. Imai, "The hierarchical architecture of nacre and its mimetic material," *Angew. Chem. Int. Ed.*, vol. 44, pp. 6571–6575, 2005.

[8] J. J. De Yoreo, P. U. P. A. Gilbert, N. A. J. M. Sommerdijk, L. R. Penn, S. Whitelam, D. Joester, H. Zhang, D. J. Rimer, A. Navrotsky, J. F. Banfield, A. F. Wallace, M. F. Michel, F. C. Meldrum, H. Colfen, and P. M. Dove, "Crystallization by particle attachment in synthetic, biogenic and geologic environments," *Science*, vol. 349, no. 6247, pp. 498–598, 2015.

[9] C. L. Freeman, J. H. Harding, D. Quigley, and M. P. Rodger, "How does an amorphous surface influence molecular binding? ovocleidin-17 and amorphous calcium carbonate," *Phys. Chem. Chem. Phys.*, vol. 17, no. 26, pp. 17494–17500, 2015.

[10] B. Hasse, H. Ehrenberg, J. C. Marxen, W. Becker, and M. Epple, "Calcium carbonate modifications in the mineralized shell of the freshwater snail Biomphalaria Glabrata," *Chemistry-a European Journal*, vol. 6, no. 20, pp. 3679–3685, 2000.

[11] Y. Levi-Kalisman, S. Raz, S. Weiner, L. Addadi, and I. Sagi, "Structural differences between biogenic amorphous calcium carbonate phases using X-ray absorption spectroscopy," *Advanced Functional Materials*, vol. 12, no. 1, pp. 43–48, 2002.

[12] L. Addadi, S. Raz, and S. Weiner, “Taking advantage of disorder: amorphous calcium carbonate and its roles in biomineralization,” *Advanced Materials*, vol. 15, no. 12, pp. 959–970, 2003.

[13] A. Becker, U. Bismayer, M. Epple, H. Fabritius, B. Hasse, J. M. Shi, and A. Ziegler, “Structural characterisation of X-ray amorphous calcium carbonate (ACC) in sternal deposits of the crustacean Porcellio Scaber,” *Dalton Transactions*, vol. 4, no. 4, pp. 551–555, 2003.

[14] Y. Levi-Kalisman, S. Raz, S. Weiner, L. Addadi, and I. Sagi, “X-ray absorption spectroscopy studies on the structure of a biogenic “amorphou” calcium carbonate phase,” *Journal of the Chemical Society-Dalton Transactions*, no. 21, pp. 3977–3982, 2000.

[15] Y. Politi, Y. Levi-Kalisman, S. Raz, F. Wilt, L. Addadi, S. Weiner, and I. Sagi, “Structural characterization of the transient amorphous calcium carbonate precursor phase in sea urchin embryos,” *Advanced Functional Materials*, vol. 16, no. 10, pp. 1289–1298, 2006.

[16] C. Gunther, A. Becker, G. Wolf, and M. Epple, “In vitro synthesis and structural characterization of amorphous calcium carbonate,” *Zeitschrift Fur Anorganische Und Allgemeine Chemie*, vol. 631, no. 13-14, pp. 2830–2835, 2005.

[17] Y. J. Wu, H. L. Tepper, and G. A. Voth, “Flexible simple point-charge water model with improved liquid-state properties,” *Journal of Chemical Physics*, vol. 124, no. 2, p. 024503, 2006.

[18] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. -, pp. 926–935, 1983.

[19] J. Gale, P. Raiteri, and R. Demichelis, "Force-field for the calcium phosphate system based on thermodynamic solution values," *Private Communication*, 2015.

[20] R. Demichelis, P. Raiteri, J. D. Gale, D. Quigley, and D. Gebauer, "Stable prenucleation mineral clusters are liquid-like ionic polymers," *Nature Communications*, vol. 2, p. 590, 2011.

[21] C. L. Freeman, J. H. Harding, D. J. Cooke, J. A. Elliott, J. S. Lardge, and D. M. Duffy, "New forcefields for modeling biomineralization processes," *J. Phys. Chem. C*, vol. 111, no. 32, pp. 11943–11951, 2007.

[22] M. Mathew, L. W. Schroeder, B. Dickens, and W. E. Brown, "The crystal structure of α-$Ca_3(PO_4)_2$," *Advanced Materials*, vol. 33, no. 5, pp. 1325–1333, 1977.

[23] M. Catti, G. Ferraris, and A. Flihol, "Hydrogen bonding in the crystalline state. $CaHPO_4$ (monetite), P1- or P1 ? a novel neutron diffraction study," *Acta Crystallographica B*, vol. 33, no. 24, pp. 1223–1229, 1977.

[24] P. A. Cohen, J. W. Schopf, N. J. Butterfield, A. B. Kudryavtsev, and F. A. Macdonald, "Phosphate biomineralization in mid-neoproterozoic protists," *Geological Society of America*, vol. 39, no. 6, pp. 539–542, 2011.

[25] A. H. Knoll, "Biomineralization and evolutionary history," *Reviews in Mineralogy and Geochemistry*, vol. 54, no. 1, pp. 329–356, 2003.

[26] W. A. Deer, R. A. Howie, and J. Zussman, *An introduction to the rock-forming minerals*. Harlow: Longman, first ed., 1992.

[27] H. Zhou and J. Lee, "Nanoscale hydroxyapatite particles for bone tissue engineering," *Acta Biomaterialia*, vol. 7, no. 7, pp. 2769–2781, 2011.

[28] W. Suchanek and M. Yoshimura, "Processing and properties of hydroxyapatite-based biomaterials for use as hard tissue replacement implants," *Journal of Materials Research*, vol. 13, no. 1, pp. 94–117, 1998.

[29] S. V. Dorozhkin and M. Epple, "Biological and medical significance of calcium phosphates," *Angewandte Chemie-International Edition*, vol. 41, no. 17, pp. 3130–3146, 2002.

[30] S. V. Dorozhkin, "Bioceramics of calcium orthophosphates," *Biomaterials*, vol. 31, no. 7, pp. 1465–1485, 2010.

[31] A. Corami, S. Mignardi, and V. Ferrini, "Cadmium removal from single- and multi-metal (Cd plus Pb plus Zn plus Cu) solutions by sorption on hydroxyapatite," *Journal of Colloid and Interface Science*, vol. 317, no. 2, pp. 402–408, 2008.

[32] R. Freitag and F. Hilbrig, "Isolation and purification of recombinant proteins, antibodies and plasmid dna with hydroxyapatite chromatography," *Biotechnology Journal*, vol. 7, no. 1, pp. 90–102, 2012.

[33] V. C. Ghantani, S. T. Lomate, M. K. Dongare, and S. B. Umbarkar, "Catalytic dehydration of lactic acid to acrylic acid using calcium hydroxyapatite catalysts," *Green Chemistry*, vol. 15, no. 5, pp. 1211–1217, 2013.

[34] N. H. de Leeuw, "Density functional theory calculations of local ordering of hydroxy groups and fluoride ions in hydroxyapatite," *Phys. Chem. Chem. Phys*, vol. 4, no. 1, pp. 3865–3871, 2002.

[35] W. Kolodziejski, "Solid-state NMR studies of bone," *Top Curr Chem*, vol. 246, no. 6, pp. 235–270, 2004.

[36] S. Peroos, Z. Du, and N. H. De Leeuw, "A computer modelling study of the uptake, structure and distribution of carbonate defects in hydroxy-apatite," *Biomaterials*, vol. 27, no. 9, pp. 2150–2161, 2005.

[37] N. H. de Leeuw, "A computer modelling study of the uptake and segregation of fluoride ions at the hydrated hydroxyapatite (0001) surface: introducing a $Ca_{10}(PO_4)_6(OH)_2$ potential model," *Phys. Chem. Chem. Phys*, vol. 6, no. 8, pp. 1860–1866, 2004.

[38] A. Knappwost, "Fluor-hydroxyl-substitution am hydroxylapatit als ionenaustauschreaktion und ihre anwendung mikrofluoranalyse," *Angewandte Chemie*, vol. 68, no. 11, pp. 371–373, 1956.

[39] N. Almora-Barrios, R. G. Crespo, and N. H. De Leeuw, "A computational study of magnesium incorporation in the bulk and surfaces of hydroxyapatite," *Langmuir*,

vol. 29, no. 19, pp. 5851–5856, 2013.

[40] F. Nudelman, A. J. Lausch, N. A. J. M. Sommerdijk, and E. D. Sone, "In vitro models of collagen biomineralization," *Journal of Structural Biology*, vol. 183, no. 2, pp. 258–269, 2013.

[41] S. Mann, *Biomineralization.* Oxford: Oxford University Press, first ed., 2001.

[42] M. J. Olszta, X. Cheng, S. S. Jee, R. Kumar, Y. Kim, M. J. Kaufman, E. P. Douglas, and L. B. Gower, "Bone structure and formation: a new perspective," *Materials Science and Engineering R-Reports*, vol. 58, no. 3-5, pp. 77–116, 2007.

[43] S. S. Jee, K. R. Kasinath, E. DiMasi, and L. B. Kim, Y. Gower, "Oriented hydroxyapatite in turkey tendon mineralized via the polymer-induced liquid-precursor (PILP) process," *CrystEngComm*, vol. 13, no. -, pp. 2077–2083, 2011.

[44] V. R. Kodati, E. G. Tomasi, J. L. Turumin, and A. T. Tu, "Raman spectroscopic identification of phosphate-type kidney stones," *Applied Spectroscopy*, vol. 45, no. 4, pp. 581–583, 1991.

[45] W. E. Brown, "Octacalcium phosphate and hydroxyapatite," *Nature*, vol. 196, no. -, pp. 1048–1050, 1962.

[46] J. Zhan, Y. H. Tseng, J. C. C. Chan, and C. Y. Mou, "Biomimetic formation of hydroxyapatite nanorods by a single-crystal-to-single-crystal transformation," *Advanced Functional Materials*, vol. 15, no. 12, pp. 2005–2010, 2005.

[47] W. Pompe, H. Worch, W. J. E. M. Habraken, P. Simon, R. Kniep, H. Ehrilich, and P. Paufler, "Octacalcium phosphate - a metastable mineral phase controls the evolution of scaffold forming proteins," *Journal of Materials Chemistry B*, vol. 3, no. -, pp. 5318–5329, 2015.

[48] N. J. Crane, V. Popescu, M. D. Morris, P. Steenhuis, and J. Ignelzi, M. A., "Raman spectroscopic evidence for octacalcium phosphate and other transient mineral species deposited during intramembranous mineralization," *Bone*, vol. 39, no. 3, pp. 434–442, 2006.

[49] D. H. Carter, P. V. Hatton, and J. E. Aaron, "The ultrastructure of slam-frozen bone mineral," *The Histochemical Journal*, vol. 29, no. 10, pp. 783–793, 1997.

[50] F. Barrere, C. M. van der Valk, R. A. J. Dalmeijer, G. Meijer, C. A. van Blitterswijk, K. de Groot, and P. Layrolle, "Osteogenecity of octacalcium phosphate coatings applied on porous metal implants," *Journal of Biomedical Materials Research Part A*, vol. 66A, no. 4, pp. 779–788, 2002.

[51] P. Habibovic, J. Li, C. M. van der Valk, G. Meijer, P. Layrolle, C. A. van Blitterswijk, and K. de Groot, "Biological performance of uncoated and octacalcium phosphate-coated Ti_6Al_4V," *Biomaterials*, vol. 26, no. 1, pp. 23–36, 2005.

[52] O. Suzuki, S. Kamakura, T. Katagiri, M. Nakamura, B. Zhao, Y. Honda, and R. Kamijo, "Bone formation enhanced by implanted octacalcium phosphate involving conversion into Ca-deficient hydroxyapatite," *Biomaterials*, vol. 27, no. 13, pp. 2671–2681, 2005.

[53] J. Mahamid, A. Sharir, L. Addadi, and S. Weiner, “Amorphous calcium phosphate is a major component of the forming fin bones of zebrafish: indications for an amorphous precursor phase,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 35, pp. 12748–12753, 2008.

[54] S. V. Dorozhkin, “Amorphous calcium (ortho)phosphates,” *Acta Biomaterialia*, vol. 6, no. 12, pp. 4457–4475, 2010.

[55] J. H. E. Cartwright, A. G. Checa, J. D. Gale, D. Gebauer, and C. I. Sainz-Diaz, “Calcium carbonate polyamorphism and its role in biomineralization: how many amorphous calcium carbonates are there?,” *Angewandte Chemie-International Edition*, vol. 51, no. 48, pp. 11960–11970, 2012.

[56] W. J. E. M. Habraken, J. Tao, L. J. Brylka, H. Friedrich, L. Bertinetti, A. S. Schenk, A. Verch, V. Dmitrovic, P. H. H. Bomans, P. M. Frederik, J. Laven, P. van der Schoot, B. Aichmayer, G. de With, J. J. DeYoreo, and N. A. J. M. Sommerdijk, “Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate,” *Nature Communications*, vol. 4, p. 1507, 2013.

[57] B. Xie, T. J. Halter, B. M. Borah, and G. H. Nancollas, “Tracking amorphous precursor formation and transformation during induction stages of nucleation,” *Crystal Growth and Design*, vol. 14, no. 4, pp. 1659–1665, 2014.

[58] J. Mahamid, B. Aichmayer, E. Shimoni, R. Ziblat, C. Li, S. Siegel, O. Paris, P. Fratzl, S. Weiner, and L. Addadi, “Mapping amorphous calcium phosphate transformation into crystalline mineral from the cell to the bone in zebrafish fin rays,” *Proceedings of*

the National Academy of Sciences of the United States of America, vol. 107, no. 14, pp. 6316–6321, 2010.

[59] J. Mahamid, A. Sharir, D. Gur, E. Zelzer, L. Addadi, and S. Weiner, “Bone mineralization proceeds through intracellular calcium phosphate loaded vesicles: a cryo-electron microscopy study,” *Journal of Structural Biology*, vol. 174, no. 3, pp. 527–535, 2011.

[60] E. N. Maslen, V. A. Strel’tsov, and N. R. Strel’tsova, “X-ray study of the electron density in calcite, $CaCO_3$,” *Acta Crystallographica*, vol. 49, no. 1, pp. 636–641, 1993.

[61] O. Karlsson and C. Lilja, “Eggshell structure, mode of development and growth rate in birds,” *Zoology*, vol. 111, no. 1, pp. 494–501, 2008.

[62] I. C. Dunn, A. B. Rodriguez-Navarro, K. Mcdade, M. Schmutz, R. Preisinger, D. Waddington, P. W. Wilson, and M. M. Bain, “Genetic variation in eggshell crystal size and orientation is large and these traits are correlated with shell thickness and are associated with eggshell matrix protein markers,” *Animal Genetics*, vol. 43, no. 4, pp. 410–418, 2011.

[63] L. Addadi, F. Nudelman, and S. Weiner, “Mollusk shell formation: a source of new concepts for understanding biomineralization processes,” *Chemistry Europe*, vol. 12, pp. 980–987, 2006.

[64] G. Wolf, J. Lerchner, H. Schmidt, H. Gamsjager, E. Konigsberger, and P. Schmidt, “Thermodynamics of $CaCO_3$ phase transitions,” *Journal of Thermal Analysis*,

vol. 46, no. 2, pp. 353–359, 1996.

[65] J. P. R. De Villiers, "Crystal structure of aragonite, strontianite and whiterite," *American Mineralogist*, vol. 56, no. 1, pp. 758–766, 1971.

[66] G. Falini, S. Albeck, S. Weiner, and L. Addadi, "Control of aragonite or calcite polymorphism by mollusk shell macromolecules," *Science*, vol. 271, no. -, pp. 67–69, 1996.

[67] B. Gotliv, L. Addadi, and S. Weiner, "Mollusk shell acidic proteins: in search of individual functions," *ChemBioChem*, vol. 4, no. 6, pp. 522–529, 2003.

[68] Y. Levi, S. Albeck, A. Brack, S. Weiner, and L. Addadi, "Control over aragonite crystal nucleation and growth: an in vitro study of biomineralization," *Chemistry Europe J.*, vol. 4, no. 3, pp. 389–396, 1998.

[69] H. J. Meyer, "Struktur und fehlordnung des vaterites," *Zeitschrift fur Kristallographie*, vol. 128, no. 3-6, pp. 183–212, 1960.

[70] K. Medeiros, E. L. Albuquerque, F. F. Maia Jr, E. W. S. Caetano, and V. N. Freire, "First-principles calculations of structural, electronic, and optical absorption properties of $CaCO_3$ vaterite," *Chemical Physical Letter*, vol. 435, no. 1, pp. 59–64, 2007.

[71] S. R. Kamhi, "On the structure of vaterite, $CaCO_3$," *Acta Cryst.*, vol. 16, no. 1, pp. 770–772, 1963.

[72] A. Le Bail, S. Ouhenia, and D. Chateigner, "Microtwinning hypothesis for a more ordered vaterite model," *Powder Diffraction*, vol. 26, no. 1, pp. 16–21, 2011.

[73] J. Wang and U. Becker, “Structure and carbonate orientation of vaterite($CaCO_3$),” *American Mineralogist*, vol. 94, no. 1, pp. 380–386, 2009.

[74] R. Demichelis, P. Raiteri, J. D. Gale, and D. Dovesi, “A new structural model for disorder in vaterite from first-principles calculations,” *CrystEngComm*, vol. 14, no. 1, pp. 44–47, 2012.

[75] U. Wehrmeister, A. L. Soldati, D. E. Jacob, T. Hager, and W. Hofmeister, “Raman spectroscopy of synthetic, geological, and biological vaterite: a Raman spectroscopic study,” *Journal of Raman Spectroscopy*, vol. 41, no. 12, pp. 193–201, 2010.

[76] R. Demichelis, P. Raiteri, J. D. Gale, and D. Dovesi, “The multiple structures of vaterite,” *Crystal Growth and Design*, vol. 13, no. 6, pp. 2247–2251, 2013.

[77] E. Mugnaioli, I. Andrusenko, T. Schuler, N. Loges, E. R. Dinnebier, M. Panthofer, W. Tremel, and U. Kolb, “*abinitio* structure determination of vaterite by automated electron diffraction,” *Angewandte Chemie Int. Ed.*, vol. 51, no. 1, pp. 7041–7045, 2012.

[78] L. Kabalah-Amitai, B. Mayzel, Y. Kauffman, A. N. Fitch, L. Bloch, P. U. P. A. Gilbert, and B. Pokroy, “Vaterite crystals contains two interspersed crystal structures,” *Science*, vol. 340, no. 1, pp. 454–457, 2013.

[79] K. M. N. Burgess and D. L. Bryce, “On the crystal structure of the vaterite polymorph of $CaCO_3$: a calcium-43 solid-state nmr and computational assessment,” *Solid State Nuclear Magnetic Resonance*, vol. 65, no. 1, pp. 75–83, 2014.

[80] J. Rodriguez-Blanco, S. Shaw, and L. G. Benning, "The kinetics and mechanisms of amorphous calcium carbonate (ACC)) crystallization to calcite, viavaterite," *Nanoscale*, vol. 3, no. 1, pp. 265–271, 2010.

[81] P. Bots, L. G. Benning, J. Rodriguez-Blanco, T. Roncal-Herrero, and S. Shaw, "Mechanistic insight into the crystallization of amorhous clcium carbonate (ACC)," *Crystal Growth and Design*, vol. 12, no. 7, pp. 3806–3814, 2012.

[82] I. P. Swainson, "The structure of monohydrocalcite and the phase composition of the beachrock deposits of Lake Butler and Lake Fellmongery, South Australia," *American Mineralogist*, vol. 93, no. 7, pp. 1014–1018, 2008.

[83] I. P. Swainson and R. P. Hammond, "Hydrogen bonding in ikaite, $CaCO_3.6(H_20)$," *Mineralogical Magazine*, vol. 67, no. 3, pp. 555–562, 2003.

[84] F. M. Michel, J. MacDonald, J. Feng, B. L. Phillips, L. Ehm, C. Tarabrella, J. B. Parise, and R. J. Reeder, "Structural characteristics of synthetic amorphous calcium carbonate," *Chemistry of Materials*, vol. 20, no. 14, pp. 4720–4728, 2008.

[85] R. J. Reeder, Y. Tang, M. P. Schmidt, L. M. Kubista, D. F. Cowan, and B. L. Phillips, "Characterization of structure in biogenic amorphous calcium carbonate: pair distribution function and nuclear magnetic resonance studies of lobster gastrolith," *Crystal Growth and Design*, vol. 13, no. 5, pp. 1905–1914, 2013.

[86] D. Gebauer, P. N. Gunawidjaja, J. Y. P. Ko, Z. Bacsik, B. Aziz, L. J. Liu, Y. F. Hu, L. Bergstrom, C. W. Tai, T. K. Sham, M. Eden, and N. Hedin, "Proto-calcite and

proto-vaterite in amorphous calcium carbonates," *Angewandte Chemie-International Edition*, vol. 49, no. 47, pp. 8889–8891, 2010.

[87] P. Raiteri and J. D. Gale, "Water is the key to nonclassical nucleation of amorphous calcium carbonate," *Journal of the American Chemical Society*, vol. 132, no. 49, pp. 17623–17634, 2010.

[88] I. M. Weiss, N. Tuross, L. Addadi, and S. Weiner, "Mollusc larval shell formation: amorphous calcium carbonate is a precursor phase for aragonite," *Journal of Experimental Zoology*, vol. 293, no. 5, pp. 478–491, 2002.

[89] Y. Politi, R. A. Metzler, M. Abrecht, B. Gilbert, F. H. Wilt, I. Sagi, L. Addadi, S. Weiner, and P. Gilbert, "Transformation mechanism of amorphous calcium carbonate into calcite in the sea urchin larval spicule," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 45, pp. 17362–17366, 2008.

[90] N. Nassif, N. Pinna, N. Gehrke, M. Antonietti, C. Jager, and H. Colfen, "Amorphous layer around aragonite platelets in nacre," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 36, pp. 12653–12655, 2005.

[91] L. B. Gower and D. J. Odom, "Deposition of calcium carbonate films by a polymer-induced liquid-precursor (pilp) process," *Journal of Crystal Growth*, vol. 210, no. 4, pp. 719–734, 2000.

[92] J. W. Gibbs, “On the equilibrium of heterogeneous substances,” *Trans. Connect. Acad. Sci.*, vol. 3, pp. 108–248, 1876.

[93] J. W. Gibbs, “On the equilibrium of heterogeneous substances,” *Trans. Connect. Acad. Sci.*, vol. 16, pp. 343–524, 1878.

[94] M. Volmer and A. Weber, “Germ-formation in oversaturated figures,” *Zeitschrift fur Physikalische Chemie*, vol. 119, no. 3/4, pp. 277–301, 1926.

[95] R. Becker and W. Doring, “Kinetische behandlung der keimbildung in ubersattigten dampfern,” *Annual Pysics*, vol. 24, no. 4, pp. 719–752, 1935.

[96] W. Pan, B. A. Kolomeisky, and P. G. Vekilov, “Nucleation of ordered solid phases of proteins via a disordered high-density state: phenomenological approach,” *The Journal of Chemical Physics*, vol. 122, no. 1, p. 174905, 2005.

[97] P. G. Vekilov, “The two-step mechanism of nucleation of crystals in solution,” *Nanoscale*, vol. 2, no. 1, pp. 2346–2357, 2010.

[98] P. R. ten Wolde and D. Frenkel, “Enhancement of protein crystal nucleation by crystal density fluctuations,” *Journal of Thermal Analysis*, vol. 46, no. 2, pp. 353–359, 1996.

[99] P. G. Vekilov, “Dense liquid precursors for the nucleation of ordered solid phases from solution,” *Crystal Growth and Design*, vol. 4, no. 4, pp. 671–685, 2004.

[100] S. Chung, Y. Kim, J. Kim, and Y. Kim, “Multiphase transformation and ostwald’s rule of stages during crystallization of a metal phosphate,” *Nature Physics*, vol. 5, no. 1, pp. 68–73, 2008.

[101] E. Loste, R. M. Wilson, R. Seshadri, and F. C. Meldrum, “The role of magnesium in stabilising amorphous calcium carbonate and controlling calcite morphologies,” *Journal of Crystal Growth*, vol. 254, no. 1-2, pp. 206–218, 2003.

[102] D. Gebauer and H. Coelfen, “Prenucleation clusters and non-classical nucleation,” *Nano Today*, vol. 6, no. 6, pp. 564–584, 2011.

[103] E. M. Pouget, P. H. H. Bomans, J. A. C. M. Goos, P. M. Frederik, G. de With, and N. A. J. M. Sommerdijk, “The initial stages of template-controlled $CaCO_3$ formation revealed by Cryo-TEM,” *Science*, vol. 323, no. 5920, pp. 1455–1458, 2009.

[104] D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom, and H. Colfen, “Pre-nucleation clusters as solute precursors in crystallisation,” *Chemical Society Reviews*, vol. 43, no. 7, pp. 2348–2371, 2014.

[105] A. F. Wallace, L. O. Hedges, A. Fernandez-Martinez, P. Raiteri, J. D. Gale, G. A. Waychunas, S. Whitelam, J. F. Banfield, and J. J. De Yoreo, “Microscopic evidence for liquid-liquid separation in supersaturated $CaCO_3$ solutions,” *Science*, vol. 341, no. 6148, pp. 885–889, 2013.

[106] M. Faatz, F. Grohn, and G. Wegner, “Amorphous calcium carbonate: synthesis and potential intermediate in biomineralization,” *Advanced Materials*, vol. 16, no. 12,

pp. 996–1000, 2004.

[107] Q. Hu, M. H. Nielsen, C. L. Freeman, L. M. Hamm, J. Tao, J. R. I. Lee, T. Y. J. Han, U. Becker, J. H. Harding, P. M. Dove, and J. J. De Yoreo, "The thermodynamics of calcite nucleation at organic interfaces: classical vs. non-classical pathways," *Faraday Discussions*, vol. 159, pp. 509–523, 2012.

[108] J. Baumgartner, A. Dey, P. H. H. Bomans, C. Le Coadou, P. Fratzl, N. A. J. M. Sommerdijk, and D. Faivre, "Nucleation and growth of magnetite from solution," *Nature Materials*, vol. 12, no. 4, pp. 310–314, 2013.

[109] A. Dey, P. H. H. Bomans, F. A. Mueller, J. Will, P. M. Frederik, G. de With, and N. A. J. M. Sommerdijk, "The role of prenucleation clusters in surface-induced calcium phosphate crystallization," *Nature Materials*, vol. 9, no. 12, pp. 1010–1014, 2010.

[110] M. L. Paine and M. L. Snead, "Tooth developmental biology: disruptions to enamel-matrix assembly and its impact on biomineralization," *Orthodontics and Craniofacial Research*, vol. 8, no. 4, pp. 239–251, 2005.

[111] M. A. Bewernitz, D. Gebauer, J. Long, H. Coelfen, and L. B. Gower, "A metastable liquid precursor phase of calcium carbonate and its interactions with polyaspartate," *Faraday Discussions*, vol. 159, pp. 291–312, 2012.

[112] L. B. Gower, "Biomimetic model systems for investigating the amorphous precursor pathway and its role in biomineralization," *Chemical Reviews*, vol. 108, no. -,

pp. 4551–4627, 2008.

[113] B. Cantaert, E. Beniash, and F. C. Meldrum, "The role of poly(aspartic acid) in the precipitation of calcium phosphate in confinement," *Journal of Materials Chemistry B*, vol. 1, no. 48, pp. 6586–6595, 2013.

[114] S. Weiner, "Aspartic acid-rich proteins:mmajor components of the soluble organic matrix of mollusk shells," *Calcified Tissue International*, vol. 1979, no. 29, pp. 163–167, 1979.

[115] B. Cantaert, Y. Kim, H. Ludwig, F. Nudelman, and F. C. Sommerdijk, N. A. J. M. Meldrum, "Think positive: phase separation enables a positively charged additive to induce dramatic changes in calcium carbonate morphology," *Advanced Functional Materials*, vol. 22, no. 5, pp. 907–915, 2012.

[116] A. V. Dobrynin, "Theory and simulations of charged polymers: from solution properties to polymeric nanomaterials," *Current Opinion in Colloid and Interface Science*, vol. 13, no. 6, pp. 376–388, 2008.

[117] A. R. Leach, "Molecular modelling," 2001.

[118] P. Sherwood, B. R. Brook, and M. S. P. Sansom, "Multiscale methods for macromolecular simulations," *Structural Biology*, vol. 18, no. 5, pp. 630–640, 2008.

[119] P. Atkins and J. De Paula, "Physical chemistry," 2002.

[120] D. Frenkel and B. Smit, "Understanding molecular simulation: from algorithms to applications," 2002.

[121] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, “Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics,” *ChemPhysChem*, vol. 6, no. 9, pp. 1809–1814, 2005.

[122] I. T. Todorov, W. Smith, K. Trachenko, and M. T. Dove, “DLPOLY3: new dimensions in molecular dynamics simulations via massive parallelism,” *Journal of Materials Chemistry*, vol. 16, no. 20, pp. 1911–1918, 2006.

[123] S. Melchionna, G. Ciccotti, and B. L. Holian, “Hoover NPT dynamics for systems varying in shape and size,” *Molecular Physics*, vol. 78, no. 3, pp. 533–544, 1993.

[124] A. F. Voter, F. Montalenti, and T. C. Germann, “Extending the time scale in atomistic simulation of materials,” *Materials Research*, vol. 32, no. 1, pp. 321–346, 2002.

[125] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, “Bond-orientational order in liquids and glasses,” *Phys. Rev. B*, vol. 28, no. 2, pp. 784–794, 1983.

[126] E. E. Santiso and B. L. Trout, “A general set of order parameters for molecular crystals,” *The Journal of Chemical Physics*, vol. 134, no. 6, pp. 1–15, 2011.

[127] P. Raiteri, R. Demichelis, and J. D. Gale, “Thermodynamically consistent force field for Molecular Dynamics simulations of alkaline-earth carbonates and their aqueous speciation,” *J. Phys. Chem. C*, vol. 119, no. 43, pp. 24447–24458, 2015.

[128] G. M. Torrie and J. P. Valleau, “Non-physical sampling distributions in Monte Carlo free-energy estimation - Umbrella Sampling,” *Journal of Computational Physics*, vol. 23, no. 2, pp. 187–199, 1977.

[129] A. S. Cote, B. Smith, and L. P. J. D., “Democritus,” 2001. Accessed: 2013-09-30.

[130] F. Betts and A. S. Posner, “X-ray radial-distribution study of amorphous calcium-phosphate,” *Materials Research Bulletin*, vol. 9, no. 3, pp. 353–360, 1974.

[131] A. S. Posner and F. Betts, “Synthetic amorphous calcium-phosphate and its relation to bone-mineral structure,” *Accounts of Chemical Research*, vol. 8, no. 8, pp. 273–281, 1975.

[132] N. Kanzaki, G. Treboux, K. Onuma, S. Tsutsumi, and A. Ito, “Calcium phosphate clusters,” *Biomaterials*, vol. 22, no. 21, pp. 2921–2929, 2001.

[133] G. Treboux, P. Layrolle, N. Kanzaki, K. Onuma, and A. Ito, “Existence of Posner’s cluster in vacuum,” *Journal of Physical Chemistry A*, vol. 104, no. 21, pp. 5111–5114, 2000.

[134] K. Onuma and A. Ito, “Cluster growth model for hydroxyapatite,” *Chemistry of Materials*, vol. 10, no. 11, pp. 3346–3351, 1998.

[135] L. Wang, S. Li, E. Ruiz-Agudo, C. V. Putnis, and A. Putnis, “Posner’s cluster revisited: direct imaging of nucleation and growth of nanoscale calcium phosphate clusters at the calcite-water interface,” *CrystEngComm*, vol. 14, no. 19, pp. 6252–6256, 2012.

[136] Q. Zhang, Y. Jiang, B. D. Gou, J. Huang, X. Y. Gao, J. T. Zhao, L. Zheng, Y. D. Zhao, and K. Wang, “In situ detection of calcium phosphate clusters in solution

and wet amorphous phase by synchrotron X-ray near-edge spectroscopy at calcium K-edge," *Crystal Growth and Design*, vol. 15, no. 5, pp. 2204–2210, 2015.

[137] S. Omelon, J. Georgiou, F. Variola, and M. N. Dean, "Colocation and role of polyphosphates and alkaline phosphatase in apatite biomineralization of elasmobranch tesserae," *Acta Biomaterialia*, vol. 10, no. 9, pp. 3899–3910, 2014.

[138] D. Zahn, "Mechanisms of calcium and phosphate ion association in aqueous solution," *Zeitschrift Fur Anorganische Und Allgemeine Chemie*, vol. 630, no. 10, pp. 1507–1511, 2004.

[139] N. Almora-Barrios and N. H. De Leeuw, "Molecular dynamics simulation of the early stages of nucleation of hydroxyapatite at a collagen template," *Crystal Growth and Design*, vol. 12, no. 2, pp. 756–763, 2012.

[140] J. Ma, "A molecular dynamics study on the nucleation of calcium phosphate regulated by collagen," *Journal of Materials Science*, vol. 49, no. 8, pp. 3099–3106, 2014.

[141] P. Raiteri, J. D. Gale, D. Quigley, and M. P. Rodger, "Derivation of an accurate force-field for simulating the growth of calcium carbonate from aqueous solution: a new model for the calcitewater interface," *J. Phys. Chem. C*, vol. 114, no. 13, pp. 5997–6010, 2010.

[142] C. Meis, J. D. Gale, L. Boyer, J. Carpena, and D. Gosset, "Theoretical study of Pu and Cs incorporation in a mono-silicate neodymium fluoroapatite %

$Ca_9Nd(Sio_4)(PO_4)_5F_2$," *The Journal of Physical Chemistry A*, vol. 104, no. 22, pp. 5380–5387, 2000.

[143] W. T. Lee, M. T. Dove, and E. K. H. Salje, "Surface relaxations in hydroxyapatite," *Journal of Physics: Condensed Matter*, vol. 48, p. 9829, 2000.

[144] D. Mkhonto and N. H. de Leeuw, "A computer modelling study of the effect of water on the surface structure and morphology of fluorapatite: introducing a $Ca_10(PO_4)_6F_2$ potential model," *The Journal of Materials Chemistry*, vol. 12, no. 9, pp. 2633–2642, 2002.

[145] S. Hauptman, H. Dufner, S. M. Brickmann, S. M. Kast, and R. S. Berry, "Potential energy function for apatites," *Physical Chemistry Chemical Physics*, vol. 3, no. 5, pp. 635–639, 2003.

[146] A. Pedone, M. Corno, B. Civalleri, G. Malavasi, M. C. Menziani, and U. Segre, "An *abinitio* parameterized interatomic force field for hydroxyapatite," *The Journal of Materials Chemistry*, vol. 20, no. 17, pp. 2061–2068, 2007.

[147] M. Catti, G. Ferraris, and M. Franchini Angela, "The crystal structure of $Na_2HPO_4(H_2O)_2$. competition between coordination and hydrogen bonds," *Acta Crystallographica B*, vol. 33, no. 24, pp. 1968–1982, 1977.

[148] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. Payne, "First principles methods using CASTEP," *Z. Kristall.*, vol. 220, pp. 567–570, 2005.

[149] J. D. Gale, “GULP: A computer program for the symmetry-adapted simulation of solids,” *Journal of the Chemical Society-Faraday Transactions*, vol. 93, no. 4, pp. 629–637, 1997.

[150] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, “PLUMED: a portable plugin for free-energy calculations with molecular dynamics,” *Computer Physics Communications*, vol. 180, no. 10, pp. 1961–1972, 2009.

[151] A. Grossfield, “WHAM: the weighted histogram analysis method,” 2.0.9.

[152] L. Martinez, R. Andrade, E. G. Birgin, and J. M. Martinez, “PACKMOL: a package for building initial configurations for Molecular Dynamics simulations,” *Journal of Computational Chemistry*, vol. 30, no. 13, pp. 2157–2164, 2009.

[153] M. S. Gordon and M. W. Schmidt, “Advances in electronic structure theory: GAMESS a decade later,” *Theory and Applications of Computational Chemistry: the First Forty Years*, pp. 1167–1189, 2005.

[154] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, “Comparison of multiple amber force fields and development of improved protein backbone parameters,” *Proteins-Structure Function and Bioinformatics*, vol. 65, no. 3, pp. 712–725, 2006.

[155] A. Chughtai, R. Marshall, and Nancolla.Gh, “Complexes in calcium phosphate solutions,” *Journal of Physical Chemistry*, vol. 72, no. 1, pp. 208–220, 1968.

[156] N. Bjerrum, *Calciumorthophosphate*. -: -, first ed., 1958.

[157] L. Du, S. Bian, B. Gou, Y. Jiang, J. Huang, Y. Gao, Y. Zhao, W. Wen, T. Zhang, and K. Wang, "Structure of clusters and formation of amorphous calcium phosphate and hydroxyapatite: from the perspective of coordination chemistry," *Crystal Growth and Design*, vol. 13, no. 7, pp. 3103–3109, 2013.

[158] Y. Marcus and G. Hefter, "Ion pairing," *Chemical Reviews*, vol. 106, no. 11, pp. 4585–4621, 2006.

[159] A. S. Deshpande and E. Beniash, "Bioinspired synthesis of mineralized collagen fibrils," *Crystal Growth and Design*, vol. 8, no. 8, pp. 3084–3090, 2008.

[160] A. S. Deshpande, P. Fang, X. Zhang, T. Jayaraman, C. Sfeir, and E. Beniash, "Primary structure and phosphorylation of dentin matrix protein 1 (DMP1) and dentin phosphophoryn (dpp) uniquely determine their role in biomineralization," *Biomacromolecules*, vol. 12, no. 8, pp. 2933–2945, 2011.

[161] M. T. Jahromi, G. Yao, and M. Cerruti, "The importance of amino acid interactions in the crystallization of hydroxyapatite," *Journal of the Royal Society Interface*, vol. 10, no. 80, 2013.

[162] S. Weiner, Y. Levi-Kalisman, S. Raz, and L. Addadi, "Biologically formed amorphous calcium carbonate," *Connective Tissue Research*, vol. 44, pp. 214–218, 2003.

[163] C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger, "Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation," *Journal of Physical Chemistry C*, vol. 115, no. 16, pp. 8175–8183, 2011.

[164] N. A. J. M. Sommerdijk and G. de With, "Biomimetic $CaCO_3$ mineralization using designer molecules and interfaces," *Chemical Reviews*, vol. 108, no. 11, pp. 4499–4550, 2008.

[165] R. S. K. Lam, J. M. Charnock, A. Lennie, and F. C. Meldrum, "Synthesis-dependant structural variations in amorphous calcium carbonate," *CrystEngComm*, vol. 9, no. 12, pp. 1226–1236, 2007.

[166] C. J. Stephens, S. F. Ladden, F. C. Meldrum, and H. K. Christenson, "Amorphous calcium carbonate is stabilized in confinement," *Advanced Functional Materials*, vol. 20, no. 13, pp. 2108–2115, 2010.

[167] E. Loste, R. J. Park, J. Warren, and F. C. Meldrum, "Precipitation of calcium carbonate in confinement," *Advanced Functional Materials*, vol. 14, no. 12, pp. 1211–1220, 2004.

[168] F. C. Meldrum, "Calcium carbonate in biomineralisation and biomimetic chemistry," *International Materials Reviews*, vol. 48, no. 3, pp. 187–224, 2003.

[169] K. Lee, W. Wagermaier, A. Masic, K. P. Kommareddy, M. Bennet, I. Manjubala, S. Lee, S. B. Park, H. Coelfen, and P. Fratzl, "Self-assembly of amorphous calcium carbonate microlens arrays," *Nature Communications*, vol. 3, 2012.

[170] A. V. Radha, T. Z. Forbes, C. E. Killian, P. U. P. A. Gilbert, and A. Navrotsky, "Transformation and crystallization energetics of synthetic and biogenic amorphous calcium carbonate," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 38, pp. 16438–16443, 2010.

[171] M. P. Schmidt, A. J. Ilott, B. L. Phillips, and R. J. Reeder, "Structural changes upon dehydration of amorphous calcium carbonate," *Crystal Growth and Design*, vol. 14, no. 3, pp. 938–951, 2014.

[172] E. Beniash, L. Addadi, and S. Weiner, "Cellular control over spicule formation in sea urchin embryos: a structural approach," *Journal of Structural Biology*, vol. 125, no. 1, pp. 50–62, 1999.

[173] D. Quigley and P. M. Rodger, "Free energy and structure of calcium carbonate nanoparticles during early stages of crystallization," *Journal of Chemical Physics*, vol. 128, no. 22, p. 221101, 2008.

[174] D. Quigley, C. L. Freeman, J. H. Harding, and P. M. Rodger, "Sampling the structure of calcium carbonate nanoparticles with metadynamics," *Journal of Chemical Physics*, vol. 134, no. 4, p. 044703, 2011.

[175] J. W. Singer, A. O. Yazaydin, R. J. Kirkpatrick, and G. M. Bowers, "Structure and transformation of amorphous calcium carbonate: a solid-state Ca-43 NMR and computational Molecular Dynamics investigation," *Chemistry of Materials*, vol. 24, no. 10, pp. 1828–1836, 2012.

[176] M. Saharay and R. J. Kirkpatrick, "Onset of orientational order in amorphous calcium carbonate (ACC) upon dehydration," *Chemical Physics Letters*, vol. 591, pp. 287–291, 2014.

[177] A. R. Finney and P. M. Rodger, "Probing the structure and stability of calcium carbonate pre-nucleation clusters," *Faraday Discussions*, vol. 159, pp. 47–60, 2012.

[178] A. Pavese, M. Catti, G. D. Price, and R. A. Jackson, "Interatomic potentials for $CaCO_3$ polymorphs (calcite and aragonite), fitted to elastic and vibrational data," *Physics and Chemistry of Minerals*, vol. 19, no. 2, pp. 80–87, 1992.

[179] Y. G. Bushuev, A. R. Finney, and P. M. Rodger, "Stability and structure of hydrated amorphous calcium carbonate," *Crystal Growth and Design*, vol. 15, no. 11, pp. 5269–5279, 2015.

[180] D. J. Price and C. L. Brooks, "A modified TIP3P water potential for simulation with Ewald summation," *Journal of Chemical Physics*, vol. 121, no. 20, pp. 10096–10103, 2004.

[181] M. L. Connolly, "Solvent-accessible surfaces of proteins and nucleic-acids," *Science*, vol. 221, no. 4612, pp. 709–713, 1983.

[182] Q. G. Zhang, Q. L. Liu, Y. Chen, J. Y. Wu, and A. M. Zhu, "Microstructure dependent diffusion of water-ethanol in swollen poly(vinyl alcohol): a Molecular Dynamics simulation study," *Chemical Engineering Science*, vol. 64, no. 2, pp. 334–340, 2009.

[183] T. R. Zeitler, J. A. Greathouse, and R. T. Cygan, "Effects of thermodynamic ensembles and mineral surfaces on interfacial water structure," *Physical Chemistry Chemical Physics*, vol. 14, no. 5, pp. 1728–1734, 2012.

[184] L. A. Rowley, D. Nicholson, and N. G. Parsonage, "Monte-carlo grand canonical ensemble calculation in a gas-liquid transition region for 12-6 Argon," *Journal of Computational Physics*, vol. 17, no. 4, pp. 401–414, 1975.

[185] D. A. McQuarrie, *Statistical thermodynamics.* Harper's chemistry series, University of Michigan: University Science Books, first ed., 1991.

[186] P. Fenter, S. Kerisit, P. Raiteri, and J. D. Gale, "Is the calcite-water interface understood? direct comparisons of molecular dynamics simulations with specular X-ray reflectivity data," *Journal of Physical Chemistry C*, vol. 117, no. 10, pp. 5028–5042, 2013.

[187] J. Bolze, B. Peng, N. Dingenouts, P. Panine, T. Narayanan, and M. Ballauff, "Formation and growth of amorphous colloidal $CaCO_3$ precursor particles as detected by time-resolved SAXS," *Langmuir*, vol. 18, no. 22, pp. 8364–8369, 2002.

[188] J. Bolze, D. Pontoni, M. Ballauff, T. Narayanan, and H. Colfen, "Time-resolved SAXS study of the effect of a double hydrophilic block-copolymer on the formation of $CaCO_3$ from a supersaturated salt solution," *Journal of Colloid and Interface Science*, vol. 277, no. 1, pp. 84–94, 2004.

[189] M. Faatz, W. Cheng, G. Wegner, G. Fytas, R. S. Penciu, and E. N. Economou, "Mechanical strength of amorphous $CaCO_3$ colloidal spheres," *Langmuir*, vol. 21, no. 15, pp. 6666–6668, 2005.

[190] G. Wolf, E. Konigsberger, H. G. Schmidt, and H. Gamsjager, "Thermodynamic aspects of the vaterite-calcite phase transition," *Journal of Thermal Analysis and Calorimetry*, vol. 60, no. 2, pp. 463–472, 2000.

[191] B. Lee and F. M. Richards, "Interpretation of protein structures - estimation of static accessibility," *Journal of Molecular Biology*, vol. 55, no. 3, pp. 379–390, 1971.

[192] F. C. Meldrum and H. Colfen, "Controlling mineral morphologies and structures in biological and synthetic systems," *Chemical Reviews*, vol. 108, no. 11, pp. 4332–4431, 2008.

[193] E. Beniash, J. Aizenberg, L. Addadi, and S. Weiner, "Amorphous calcium carbonate transforms into calcite during sea urchin larval spicule growth," *Proc. R. Soc. Lond. B*, vol. 264, no. 1380, pp. 461–465, 1997.

[194] A. M. Belcher, X. H. Wu, R. J. Christensen, G. D. Stucky, and D. E. Morse, "Control of crystal phase switching and orientation by soluble mollusc-shell proteins," *Nature*, vol. 381, no. -, pp. 56–58, 1996.

[195] J. S. Evans, "tuning in to mollusk shell nacre- and prismatic-associated protein terminal sequences. implications for biomineralization and the construction of high per-

formance inorganic-organic composites," *Chem. Rev.*, vol. 108, no. 11, pp. 4455–4462, 2008.

[196] J. Aizenberg, G. Lambert, L. Addadi, and S. Weiner, "Stabilization of amorphous calcium carbonate by specialized macromolecules in biological and synthetic precipitates," *Advanced Materials*, vol. 8, no. 3, pp. 222–226, 1996.

[197] J. Aizenberg, G. Lambert, S. Weiner, and L. Addadi, "Factors involved in the formation of amorphous and crystalline calcium carbonate: a study of an ascidian skeleton," *Chem. Rev.*, vol. 124, no. 1, pp. 32–39, 2002.

[198] C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger, "Structural control of crystal nuclei by an eggshell protein," *Angewandte Chemie Int.*, vol. 49, no. 30, pp. 5135–5137, 2010.

[199] C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger, "Protein binding on stepped calcite surfaces: simulations of ovocleidin-17 on calcite 31.16 and 31.8," *Phys. Chem. Chem. Phys.*, vol. 14, no. 20, pp. 7287–7295, 2012.

[200] P. Raiteri, D. Demichelis, J. D. Gale, M. Kellermeier, D. Gebauer, D. Quigley, L. B. Wright, and T. R. Walsh, "Exploring the influence of organic species on pre- and post-nucleation calcium carbonate," *Faraday Discussions*, vol. 159, no. 3, pp. 61–85, 2012.

[201] M. Saharay and R. J. Kirkpatrick, "*abinitio* and metadynamics studies on the role of essential functional groups in biomineralization of calcium carbonate and environ-

mental situations," *Phys. Chem. Chem. Phys.*, vol. 16, pp. 26843–26854, 2014.

[202] M. E. Hodson, L. G. Benning, B. Demarchi, K. E. H. Penkman, J. D. Rodriguez-Blanco, P. F. Schofield, and E. A. A. Versteegh, "Biomineralisation by earthworms an investigation into the stability and distribution of amorphous calcium carbonate," *Geochemical Transactions*, vol. 16, no. 4, pp. 1–16, 2015.

[203] R. Innocenti Malini, Y. G. Bushuev, S. A. Hall, C. L. Freeman, M. P. Rodger, and J. H. Harding, "Using simulation to understand the structure and properties of hydrated amorphous calcium carbonate," *CrystEngComm*, vol. 18, no. 92, pp. 92–101, 2016.

[204] T. W. Smith, T. R. Forester, and I. T. Todorov, "DL_POLY Classic 1.9 manual," pp. 1–326, 2010.

[205] J. Kahlen, L. Salimi, M. Sulpizi, C. Peter, and D. Donadion, "Interaction of charged amino-acid side chains with ions: an optimization strategy for classical force fields," *The Journal of Physical Chemistry B*, vol. 118, no. 14, p. 39603972, 2014.

[206] S. Piana, J. L. Klepeis, and D. E. Shaw, "Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations," *Current Opinion in Structural Biology*, vol. 24, no. 1, pp. 98–105, 2014.

[207] D. J. Sparks, M. E. Romero-Gonzalez, E. El-Taboni, C. L. Freeman, S. A. Hall, G. Kakonyi, S. A. Swanson, L. Banwart, and J. H. Harding, "Adsorption of poly

acrylic acid onto the surface of calcite: an experimental and simulation study," *Phys. Chem. Chem. Phys.*, vol. 17, no. 41, pp. 27357–27365, 2015.

[208] A. S. Cote, R. Darkins, and D. M. Duffy, "Deformation twinning and the role of amino acids and magnesium in calcite hardness from molecular simulation," *Phys. Chem. Chem. Phys.*, vol. 17, no. 31, pp. 20178–20184, 2015.

[209] J. R. Dorvee and A. Veis, "Water in the formation of biogenic minerals: peeling away the hydration layers," *Journal of Structural Biology*, vol. 183, no. 1, pp. 278–303, 2013.

[210] S. Kerisit, D. J. Cooke, D. Spagnoli, and S. C. Parker, "Molecular Dynamics simulations of the interactions between water and inorganic solids," *Journal of Materials Chemistry*, vol. 15, no. 14, pp. 1454–1462, 2005.

[211] S. Kerisit and C. Liu, "Molecular simulation of the diffusion of uranyl carbonate species in aqueous solution," *Geochimica et Cosmochimica Acta*, vol. 74, no. 17, pp. 4937–4952, 2010.

[212] M. D. Fayer, "Dynamics of water interacting with interfaces, molecules, and ions," *Advanced Materials*, vol. 8, no. 3, pp. 222–226, 2011.

[213] P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shuray, and B. Lentz, ""Squiggle-H_2O. an enquiry into the importance of solvation effects in phosphate ester and anhydride reactions.," *BBA - Bioenergetics*, vol. 223, no. 1, pp. 1–15, 1970.

[214] J. Florian and A. Warshel, "Langevin dipoles model for *ab initio* calculations of chemical processes in solution: parametrization and application to hydration free energies of neutral and ionic solutes and conformational analysis in aqueous solution.," *J. Phys. Chem. B.*, vol. 101, no. 28, pp. 5583–5595, 1997.

[215] P. E. Mason, J. M. Cruickshank, G. W. Neilson, and P. Buchanan, "Neutron scattering studies on the hydration of phosphate ions in aqueous solutions of K_3PO_4, K_2HPO_4 and KH_2PO_4," *Phys. Chem. Chem. Phys.*, vol. 5, no. 20, pp. 4686–4690, 2003.

[216] L. Liang, P. Rulis, and W. Y. Ching, "Mechanical properties, electronic structure and bonding of α- and β-tricalcium phosphates with surface characterization," *Acta Biomaterialia*, vol. 6, no. 9, pp. 3763–3771, 2010.